



TUGAS AKHIR - SS 141501

**PEMODELAN *STATISTICAL DOWNSCALING*
DENGAN *PROJECTION PURSUIT REGRESSION*
UNTUK MERAMALKAN CURAH HUJAN BULANAN
DI SENTRA PRODUKSI PADI JAWA TIMUR**

**VELLA ROCHMANA ASYEIFA
NRP 1313 100 073**

**Dosen Pembimbing
Dr. Sutikno, S.Si, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS141501

**PEMODELAN *STATISTICAL DOWNSCALING*
DENGAN *PROJECTION PURSUIT REGRESSION*
UNTUK MERAMALKAN CURAH HUJAN
BULANAN DI SENTRA PRODUKSI PADI
JAWA TIMUR**

**VELLA ROCHMANA ASYEIFA
NRP 1313 100 073**

**Dosen Pembimbing
Dr. Sutikno, S.Si, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS141501

**STATISTICAL DOWNSCALING MODELLING
WITH PROJECTION PURSUIT REGRESSION
TO FORECAST MONTHLY RAINFALL
IN EAST JAVA RICE PRODUCTION CENTER**

**VELLA ROCHMANA ASYEIFA
NRP 1313 100 073**

**Supervisor
Dr. Sutikno, S.Si, M.Si**

**UNDERGRADUATE PROGRAM
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY
SURABAYA 2017**

LEMBAR PENGESAHAN

PEMODELAN *STATISTICAL DOWNSCALING* DENGAN *PROJECTION PURSUIT REGRESSION* UNTUK MERAMALKAN CURAH HUJAN BULANAN DI SENTRA PRODUKSI PADI JAWA TIMUR

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh:

Vella Rochmana Asyeifa
NRP. 1313 100 073

Disetujui oleh Pembimbing:

Dr. Sutikno, S.Si, M.Si

NIP. 19710313 199702 1 001

()



Mengetahui,
Kepala Departemen



Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

(Halaman ini sengaja dikosongkan)

PEMODELAN *STATISTICAL DOWNSCALING* DENGAN *PROJECTION PURSUIT REGRESSION* UNTUK MERAMALKAN CURAH HUJAN BULANAN DI SENTRA PRODUKSI PADI JAWA TIMUR

Nama : Vella Rochmana Asyeifa
NRP : 1313 100 073
Departemen : Statistika
Dosen Pembimbing : Dr. Sutikno, S.Si, M.Si

Abstrak

Peramalan curah hujan di wilayah yang memiliki iklim tropis, seperti Indonesia, memiliki permasalahan yang kompleks. Topografi dan interaksi antara lautan, daratan dan atmosfer yang kompleks mempersulit prediksi curah hujan di Indonesia. Oleh karena itu, diperlukan model peramalan curah hujan yang akurat pada skala lokal dengan mempertimbangkan informasi tentang sirkulasi atmosfer global yang didapatkan dari luaran General Circulation Model (GCM). GCM bisa digunakan untuk memperoleh informasi dalam skala lokal dengan melakukan statistical downscaling. Statistical downscaling merupakan model basis regresi untuk menentukan hubungan fungsional antara variabel respon dan variabel prediktor. Variabel respon yang digunakan dalam penelitian ini adalah observasi curah hujan, sedangkan variabel prediktor adalah iklim global luaran dari GCM. Penelitian ini dilakukan di sentra produksi padi Jawa Timur, yaitu Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi. Metode statistical downscaling yang digunakan adalah Projection Pursuit Regression (PPR). Dalam pemodelan PPR optimalisasi dilakukan sebanyak lima kali dengan simulasi banyak fungsi 1 sampai 5. Validasi model dilakukan dengan kriteria Root Mean Square Error Prediction (RMSEP) terkecil. Pola antara hasil ramalan dan observasi menunjukkan bahwa hasil ramalan curah hujan di lima kabupaten mendekati data observasi. Sehingga model yang diperoleh merupakan model yang baik untuk meramalkan curah hujan di lima kabupaten.

Kata Kunci: *GCM, peramalan curah hujan, PPR, statistical downscaling*

(Halaman ini sengaja dikosongkan)

STATISTICAL DOWNSCALING MODELLING WITH PROJECTION PURSUIT REGRESSION TO FORECAST MONTHLY RAINFALL IN EAST JAVA RICE PRODUCTION CENTER

Name : Vella Rochmana Asyeifa
NRP : 1313 100 073
Department : Statistics
Supervisor : Dr. Sutikno, S.Si, M.Si

Abstract

The prediction of rainfall in the area with the climate tropical, such as Indonesia, is hard to do. The Topography and interaction complex between sea, mainland, and atmosphere makes the prediction of rainfall in Indonesia complicated. Therefore, the prediction models precipitation accurate on the local scale consider the information about the circulation global atmosphere established from General Circulation Model (GCM) are required. GCM can be used to acquire the information in the scale of the local or regional by doing statistical downscaling. Statistical downscaling is a regression based model to determine the functional relationship between variable response and variable predictors. In this research variable response uses data of rainfall observation, meanwhile variable predictor uses climate scale global from GCM. The research of rainfall take place in rice production center in East Java, included Ngawi, Lamongan, Bojonegoro, Jember and Banyuwangi regency. The exact statistical downscaling method in this research is Projection Pursuit Regression (PPR). PPR modeling have five times optimizing with simulation function 1 until 5. Validation model criteria use Root Mean Square Error Prediction (RMSEP). The best model choose based on how many function which has the smallest RMSEP value. The pattern between forecasting and observation shows the result of rainfall forecasting in five regency are close to the data observation. So the obtained model is a good model to forecast rainfall in five regency.

Keywords: *GCM, PPR, rainfall forecasting, statistical downscaling*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, atas limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan laporan Tugas Akhir yang berjudul **“PEMODELAN *STATISTICAL DOWNSCALING* DENGAN *PROJECTION PURSUIT REGRESSION* UNTUK MERAMALKAN CURAH HUJAN BULANAN DI SENTRA PRODUKSI PADI JAWA TIMUR”**

Keberhasilan penulis dalam penyusunan Tugas Akhir ini tidak lepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Dr. Suhartono selaku Kepala Departemen Statistika ITS.
2. Dr. Sutikno, S.Si, M.Si selaku dosen pembimbing Tugas Akhir yang telah membantu selama proses pelaksanaan Tugas Akhir.
3. Dr.rer.pol. Heri Kuswanto, S.Si, M.Si dan Dr. Santi Wulan Purnami, S.Si, M.Si selaku dosen penguji yang telah memberikan kritik dan saran dalam penyusunan Tugas Akhir.
4. Dr. Irhamah, M.Si selaku dosen wali yang sudah memberikan nasehat, dukungan dan masukan selama masa perkuliahan.
5. Bapak Syaifudin Zuhri, Ibu Sri Setyowati dan Adik Nafakhatur Ridwan Asyeifi yang telah memberikan do'a, dukungan dan motivasi kepada penulis untuk menyelesaikan Tugas Akhir
6. Semua pihak yang tidak dapat disebutkan satu-persatu yang telah membantu hingga penyusunan Tugas Akhir ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa dalam penulisan laporan Tugas Akhir ini masih banyak kekurangan baik pada teknik penulisan maupun materi. Oleh karena itu, penulis memohon maaf serta mengharapkan kritik dan saran dari semua pihak demi penyempurnaan laporan ini.

Surabaya, Juli 2017
Penulis

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii

BAB I PENDAHULUAN

1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan	5
1.4 Manfaat	5
1.5 Batasan Masalah	6

BAB II TINJAUAN PUSTAKA

2.1 <i>General Circulation Model</i> (GCM)	7
2.2 <i>Statistical Downscaling</i>	9
2.3 <i>Projection Pursuit</i> (PP).....	11
2.4 <i>Projection Pursuit Regression</i> (PPR)	12
2.5 Validasi Model.....	17
2.6 Curah Hujan.....	18
2.7 Penelitian Sebelumnya.....	20

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data	23
3.2 Variabel Penelitian.....	23
3.3 Struktur Data.....	25
3.4 Langkah Analisis Data.....	25

BAB IV ANALISIS DAN PEMBAHASAN

- 4.1 Deskripsi Curah Hujan di Sentra Produksi Padi Jawa Timur.....29
- 4.2 Pemodelan *Statistical Downscaling* dengan *Projection Pursuit Regression* di Sentra Produksi Padi Jawa Timur.34

BAB V KESIMPULAN DAN SARAN

- 5.1 Kesimpulan43
- 5.2 Saran44

DAFTAR PUSTAKA45

LAMPIRAN47

DAFTAR TABEL

	Halaman
Tabel 2.1 Derajat Hujan Berdasarkan Intensitas Curah Hujannya	19
Tabel 3.1 Lokasi Stasiun Pengamatan.....	23
Tabel 3.2 Variabel GCM dari Aplikasi Model NCEP CFSR	24
Tabel 3.3 Struktur Data Penelitian	25
Tabel 4.1 Nilai RMSEP dan R^2_{pred} Menurut Banyaknya Fungsi di Lima Kabupaten.....	35
Tabel 4.2 Model PPR Data Curah Hujan di Lima Kabupaten.....	39

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Penentuan Pemilihan <i>Grid</i> di Setiap Lokasi Pengamatan (BMKG, 2011)	24
Gambar 3.2 Diagram Alir Pemodelan <i>Statistical Downscaling</i> dengan PPR.....	26
Gambar 4.1 <i>Boxplot</i> Curah Hujan di Wilayah Pengamatan	30
Gambar 4.2 <i>Time Series Plot</i> Curah Hujan di Wilayah Pengamatan.....	31
Gambar 4.3 <i>Scatterplot</i> Antara Variabel Respon dan Variabel Prediktor di Kabupaten Ngawi	32
Gambar 4.4 Nilai RMSEP Menurut Banyaknya Fungsi di Lima Kabupaten	36
Gambar 4.5 Nilai R^2_{pred} Menurut Banyaknya Fungsi di Lima Kabupaten	36
Gambar 4.6 <i>Plot</i> Antara Observasi dan Ramalan Curah Hujan di Kabupaten Ngawi	40
Gambar 4.7 <i>Plot</i> Antara Observasi dan Ramalan Curah Hujan di Kabupaten Lamongan	40
Gambar 4.8 <i>Plot</i> Antara Observasi dan Ramalan Curah Hujan di Kabupaten Bojonegoro	41
Gambar 4.9 <i>Plot</i> Antara Observasi dan Ramalan Curah Hujan di Kabupaten Jember	41
Gambar 4.10 <i>Plot</i> Antara Observasi dan Ramalan Curah Hujan di Kabupaten Banyuwangi	42

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran A. Statistika Deskriptif Curah Hujan Berdasarkan Bulan	47
Lampiran B. <i>Scatterplot</i> Antara Variabel Respon dan Variabel Prediktor.....	50
Lampiran C. Matriks Koefisien Proyeksi α_{k1}	58
Lampiran D. <i>Syntax Projection Pursuit Regression</i> (PPR)	61
Lampiran E. Surat Keterangan Legalisasi Data	62

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Indonesia merupakan suatu negara yang dilewati oleh garis khatulistiwa dan mempunyai daerah perairan dengan luasan relatif besar. Hal tersebut mengakibatkan wilayah Indonesia memiliki karakteristik yang berbeda dibandingkan dengan atmosfer di daerah khatulistiwa lainnya. Indonesia juga dikenal dengan sebutan Benua Maritim Indonesia atau *Indonesia Maritime Continent* (IMC) karena letak geografisnya yang berada di antara Benua Asia dan Benua Australia serta di antara Samudera Pasifik dan Samudera Hindia. Secara umum cuaca dan iklim di Indonesia dipengaruhi oleh beberapa fenomena alam, antara lain gelombang El-Nino, monsun Asia-Australia, sirkulasi atmosfer lokal, regional dan global, seperti sirkulasi utara-selatan, sirkulasi barat-timur dan sistem angin lokal. Apabila salah satu dari sistem sirkulasi tersebut terganggu, maka secara otomatis akan mempengaruhi cuaca dan iklim di Indonesia (Mukarami dan Zadrach, 2000).

Indonesia memiliki ciri iklim yang khas yaitu keragaman curah hujan yang cukup besar antar daerah di Indonesia. Iklim dapat didefinisikan sebagai kebiasaan cuaca yang terjadi di suatu tempat atau daerah yang umumnya muncul secara periodik atau berkala dan terjadi dalam periode waktu yang lama (Sucahyono dan Ribudiyanto, 2013). Iklim memiliki dua unsur penting, yaitu suhu dan curah hujan. Indonesia sebagai negara beriklim tropis yang dilewati oleh garis khatulistiwa memiliki keragaman suhu yang kecil, tetapi memiliki keragaman curah hujan yang sangat besar. Oleh karena itu, penelitian terhadap curah hujan lebih sering dilakukan daripada melakukan penelitian terhadap unsur iklim lainnya yaitu suhu.

Badan Meteorologi, Klimatologi dan Geofisika (BMKG) sebagai pusat peramalan cuaca dan iklim di Indonesia telah membuat peramalan curah hujan untuk kota-kota besar yang ada di Indonesia. Namun, peramalan curah hujan secara terperinci sampai

ke daerah-daerah masih sangat sedikit. Peramalan curah hujan di wilayah yang memiliki iklim tropis, termasuk Indonesia, merupakan sesuatu yang sangat sulit dilakukan. Topografi dan interaksi antara lautan, daratan dan atmosfer yang kompleks mempersulit prediksi curah hujan di Indonesia. Oleh karena itu, model peramalan curah hujan harus terus ditingkatkan, salah satunya adalah dengan membuat model peramalan curah hujan yang akurat pada skala lokal dengan mempertimbangkan informasi tentang sirkulasi atmosfer global yang didapatkan dari luaran *General Circulation Model* (GCM) (Wilby dan Wigley, 1997).

GCM merupakan model numerik yang bersifat deterministik dengan simulasi komputer yang rumit yang mampu menggambarkan kondisi iklim dengan berbagai macam komponennya yang berubah sepanjang waktu (Wigena, 2006). GCM mampu mensimulasikan variabel-variabel iklim global pada setiap *grid* (berukuran $\pm 300 \text{ km}^2$) di setiap lapisan atmosfer, yang dapat digunakan untuk memprediksi pola iklim dalam jangka waktu yang panjang, yaitu jangka waktu tahunan. Namun, informasi yang dikeluarkan GCM masih mempunyai skala global dan tidak untuk fenomena dengan skala yang lebih kecil, sehingga sulit untuk menerima informasi berskala lokal secara langsung dari luaran GCM. GCM memiliki resolusi yang terlalu rendah (tidak bisa lebih kecil dari 100 km^2) untuk memprediksi iklim lokal. Akan tetapi, GCM masih bisa digunakan untuk memperoleh informasi dalam skala lokal atau regional dengan melakukan teknik *downscaling* (Fernandez, 2005).

Dalam melakukan teknik *downscaling* iklim regional dan lokal berada pada kondisi iklim skala global, dimana informasi pada skala kecil diperoleh dari informasi pada skala besar (Storch, Hewitson dan Mearns, 2001 dalam Wigena 2006). Pada penelitian ini teknik *downscaling* yang digunakan adalah *statistical downscaling*. Pendekatan *statistical downscaling* dinyatakan dengan model statistik yaitu model regresi untuk menentukan hubungan fungsional antara variabel respon yang berupa iklim lokal dan variabel prediktor yang berupa iklim skala global luaran

dari GCM. Namun, data GCM tidak dapat digunakan secara langsung sebagai variabel prediktor pada model regresi biasa. Data luaran GCM secara umum merupakan data spasial dan temporal yang memiliki kemungkinan besar untuk terjadi korelasi spasial antar data pada *grid* yang berbeda dalam satu domain dan otokorelasi pada dalam deret waktu. Di samping itu domain GCM merupakan salah satu faktor dalam *statistical downscaling*, yaitu faktor lokasi dan luasan area permukaan atmosfer. Dalam pemodelan *statistical downscaling* domain adalah sebagai variabel prediktor yang berdimensi banyak yang besar kemungkinan terjadi *curse of dimensionality*, korelasi spasial antar *grid* dalam domain dan multikolinearitas antar variabel prediktor. Semakin besar domain dan semakin banyak variabel yang digunakan akan mengakibatkan model semakin kompleks.

Masalah yang muncul dalam pemodelan *statistical downscaling* adalah pereduksian domain GCM dan penyusunan model hubungan antara variabel respon dan variabel prediktor. Metode yang seringkali digunakan dalam mereduksi dimensi tinggi adalah analisis komponen utama (*Principal Component Analysis* - PCA). Selain itu, terdapat metode lain yang juga dapat digunakan untuk mereduksi dimensi tinggi, yaitu *Projection Pursuit* (PP). Kelebihan yang dimiliki oleh metode PP dibandingkan metode PCA adalah hubungan antara variabel asal dan variabel baru dapat disusun secara nonlinier, sedangkan pada PCA hubungan tersebut harus linier (Friedman dan Stuetzle, 1981). Selain itu, PPR juga dapat mengatasi multikolinearitas antar variabel prediktor dan sifat variabel respon yang nonlinier serta tidak berdistribusi baku (misalnya distribusi normal). Oleh karena itu, metode PP lebih tepat digunakan dalam mereduksi data GCM. Adapun hasil reduksi dimensi PP kemudian digunakan untuk melakukan pemodelan PPR.

Penelitian yang dilakukan menggunakan metode *statistical downscaling* telah banyak dilakukan dalam bidang meteorologi di Indonesia, antara lain Wigena (2006) yang melakukan penelitian mengenai prediksi curah hujan bulanan di Indramayu

menggunakan pemodelan *statistical downscaling* dengan pendekatan *Projection Pursuit Regression* (PPR). Selain itu Anitawati (2010) juga pernah melakukan penelitian mengenai pemodelan *statistical downscaling* dengan metode *Principal Component Regression* (PCR) dan *Projection Pursuit regression* (PPR), dimana hasil penelitian tersebut menunjukkan bahwa pemodelan dengan metode PPR memiliki hasil yang lebih baik dibandingkan dengan metode PCR. Metode PPR tidak hanya dilakukan untuk melakukan peramalan cuaca jangka panjang tetapi juga digunakan untuk meramalkan cuaca jangka pendek. Salah satunya adalah penelitian yang dilakukan oleh Safitri (2012) yaitu meramalkan suhu minimum, suhu maksimum dan kelembaban menggunakan *Model Output Statistics* (MOS) dengan metode PPR. Berdasarkan berbagai penelitian tersebut metode PPR menghasilkan model yang baik karena PPR dapat mengatasi berbagai permasalahan dalam melakukan peramalan cuaca, seperti multikolinearitas antar variabel prediktor dan sifat variabel respon yang nonlinier serta tidak berdistribusi baku.

Kondisi iklim selalu menjadi topik yang hangat, baik di Indonesia maupun di seluruh dunia. Hal tersebut dikarenakan salah satu unsur iklim yaitu curah hujan memiliki keragaman yang sangat tinggi. Peramalan curah hujan di Indonesia sangat sulit dilakukan dikarenakan topografi dan interaksi antara lautan, daratan dan atmosfer yang kompleks. Oleh karena itu, diperlukan suatu model peramalan curah hujan yang akurat pada skala lokal. Model yang tepat digunakan untuk saat ini adalah model *statistical downscaling* dengan metode PPR dikarenakan metode tersebut mampu mengatasi berbagai permasalahan dalam melakukan pemodelan curah hujan. Adapun penelitian curah hujan dilakukan di sentra produksi padi Jawa Timur, yaitu Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi. Lokasi tersebut dipilih karena informasi iklim sangat berguna untuk meningkatkan produksi dan kualitas padi sertaantisipasi pergeseran awal musim dan periode masa tanam (Runtunuwu dan Syahbuddin, 2007). Pemodelan tersebut diharapkan dapat menghasilkan peramalan

curah hujan yang lebih baik dibandingkan dengan metode lainnya serta dapat memberikan informasi terhadap instansi terkait dalam melakukan peramalan curah hujan untuk jangka waktu yang panjang. Peramalan curah hujan tersebut juga dapat dimanfaatkan dalam bidang pertanian khususnya untuk menyusun kalender tanam padi.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian di atas, permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut.

1. Bagaimana menyusun model *statistical downscaling* dengan metode PPR untuk meramalkan curah hujan bulanan di sentra produksi padi Jawa Timur.
2. Bagaimana keakuratan hasil ramalan model *statistical downscaling* dengan metode PPR.

1.3 Tujuan

Berdasarkan rumusan masalah tersebut, tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Mendapatkan model *statistical downscaling* dengan metode PPR untuk meramalkan curah hujan bulanan di sentra produksi padi Jawa Timur.
2. Mengetahui keakuratan hasil ramalan model *statistical downscaling* dengan metode PPR.

1.4 Manfaat

Manfaat yang diharapkan melalui penelitian ini adalah model *statistical downscaling* diharapkan bisa menjadi alternatif baru bagi BMKG maupun institusi lainnya dalam melakukan peramalan cuaca jangka panjang serta dapat membantu Depatemen Pertanian dengan memberikan informasi iklim yang digunakan untuk menyusun kalender tanam padi. Model *statistical downscaling* yang digunakan dalam penelitian ini dapat digunakan sebagai upaya untuk mendeteksi peristiwa-peristiwa perubahan cuaca dalam jangka panjang. Selain itu, penelitian ini diharapkan dapat menambah pengetahuan mengenai metode PPR yang

digunakan dalam melakukan peramalan cuaca jangka panjang dengan menggunakan model *statistical downscaling*.

1.5 Batasan Masalah

Berdasarkan latar belakang dan rumusan masalah di atas, batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut.

1. GCM yang digunakan dalam penelitian ini adalah hasil model aplikasi NCEP CFSR (*National Centers for Environmental Prediction - Climate Forecast System Reanalysis*) dengan resolusi sebesar $0,3^0 \times 0,3^0$ yang kurang lebih setara dengan $35 \text{ km} \times 35 \text{ km}$.
2. Unsur iklim yang digunakan dalam penelitian ini adalah curah hujan.

BAB II

TINJAUAN PUSTAKA

2.1 *General Circulation Model (GCM)*

GCM merupakan model numerik yang bersifat deterministik dengan simulasi komputer yang rumit yang dapat menggambarkan kondisi iklim dengan berbagai macam komponennya yang berubah sepanjang waktu. Model ini menggambarkan hubungan matematik dari berbagai interaksi fisika, kimia dan dinamika atmosfer bumi. Model ini merupakan model yang mampu memahami iklim di masa lampau, sekarang dan masa mendatang. GCM mampu mensimulasikan variabel-variabel iklim global pada setiap *grid* (berukuran $\pm 300 \text{ km}^2$) di setiap lapisan atmosfer, yang dapat digunakan untuk memprediksi pola iklim dalam jangka waktu yang panjang, yaitu jangka waktu tahunan.

GCM merupakan suatu alat yang penting dalam penelitian keragaman dan perubahan iklim (Zorita dan Storch, 1999). Model ini mampu menggambarkan berbagai macam subsistem iklim di bumi, seperti proses-proses di lautan, daratan dan atmosfer. Selain itu GCM juga mampu mensimulasi kondisi iklim berskala besar atau beresolusi rendah, misalnya dapat memproduksi dengan baik pola-pola keragaman atmosfer dan temperatur permukaan laut. Namun, GCM tidak dirancang untuk menghasilkan informasi penting dengan resolusi yang lebih tinggi atau skala yang lebih kecil, misalnya untuk temperatur dan curah hujan skala lokal. Dengan demikian, GCM dapat melakukan simulasi dengan baik untuk variabel iklim skala besar, tetapi tidak untuk variabel dengan skala yang lebih kecil (lokal atau regional). Menurut Zorita dan Storch (1999) alasan mengapa GCM tidak dapat menghasilkan informasi untuk skala lokal atau regional adalah sebagai berikut.

1. Deskripsi solusi spasial tentang struktur permukaan bumi tidak jelas, khususnya topografi.
2. Hidrodinamika atmosfer bersifat nonlinier sehingga terdapat interaksi nonlinier antara *grid* skala kecil.

3. Terlalu banyak parameter yang tidak mungkin tepat digunakan untuk proses-proses pada skala kecil.

GCM perlu disempurnakan secara terus menerus dan bertahap, meskipun model ini belum memadai untuk fenomena berskala kecil di kawasan topografi yang kompleks. Oleh karena itu, GCM tidak dapat digunakan secara langsung untuk melakukan penelitian terhadap dampak iklim secara regional maupun lokal. Adapun langkah-langkah yang dilakukan dalam melakukan penyempurnaan terhadap GCM untuk mengatasi perbedaan skala adalah meningkatkan kemampuan komputer terutama untuk skala yang lebih kecil, sehingga dapat melakukan *Regional Circulation Model* (RCM) dari luaran GCM. Disamping itu juga perlu menerapkan teknik *statistical downscaling*.

Pada umumnya, GCM telah digunakan dan dikembangkan sejak lama di sebagian besar pusat peramalan cuaca dan iklim utama di dunia. Seperti *Goddard Institute for Space Studies* (GISS) dari NASA, *United Kingdom Meteorological Office* (UKMO), *Geophysical Fluid Dynamic Laboratory* (GFDL) dari NOAA, *National Centers for Environmental Prediction* (NCEP) dari Amerika, *Commonwealth Scientific and Industrial Research Organization* (CSIRO) dari Australia, dan sebagainya. Masing-masing dari GCM tersebut memiliki perbedaan dalam bentuk resolusi spasial dan persamaan-persamaan untuk membangkitkan parameter-parameter atmosfer. GCM juga telah digunakan dan dikembangkan di Indonesia untuk melakukan simulasi, prediksi dan pembuatan skenario iklim (Ratag 2001).

Pemodelan *statistical downscaling* memerlukan domain GCM yang terdiri dari sejumlah *grid* yang berada pada suatu lokasi tertentu. Data luaran GCM dalam suatu domain secara umum bersifat *curse of dimensionality*. Hal tersebut seringkali menjadi masalah apabila dimensi atau domain semakin besar, yaitu jika pemodelan *statistical downscaling* melibatkan banyak variabel dan berbagai macam lapisan atmosfer. Dalam hal ini data bersifat nonlinier dan tidak berdistribusi yang baku, seperti distribusi normal. Di samping itu masalah lain yang berkaitan dengan data

adalah terjadinya korelasi spasial dan/atau multikolinieritas antar variabel. GCM menghasilkan luaran untuk berbagai variabel pada berbagai lapisan atmosfer atau ketinggian. Semakin banyak jumlah variabel dan lapisan atmosfer yang digunakan dalam pemodelan *statistical downscaling*, permasalahan data juga akan semakin kompleks.

2.2 *Statistical Downscaling*

GCM memiliki resolusi yang rendah dan tidak memiliki resolusi yang lebih kecil dari 100 km², sehingga model ini tidak dapat meresolusikan kejadian-kejadian atau efek-efek yang berskala lokal dari keberadaan heterogenitas topografi, vegetasi, dan komposisi tanah (Ratag, 2001). Parameter-parameter skala kecil GCM dapat digunakan untuk memprediksi variabilitas dan perubahan lokal dengan syarat wilayah tersebut memiliki topografi yang relatif homogen, seperti pada wilayah-wilayah dengan perubahan ketinggian yang relatif kecil. Perubahan iklim yang disimulasikan terjadi pada satu *grid* kemungkinan akan berlaku secara merata pada wilayah-wilayah homogen tersebut. Sebaliknya, pada kawasan dengan variasi topografi yang besar sangat dipengaruhi oleh iklim regional, sehingga untuk memprediksi variabel lokal diperlukan pendekatan *downscaling*.

Terdapat dua jenis pendekatan *downscaling*, yaitu *dynamical downscaling* dan *statistical downscaling*. *Dynamical downscaling* dilakukan dengan cara menetapkan model GCM tersarang dengan resolusi spasial yang lebih tinggi, sedangkan *statistical downscaling* dilakukan berdasarkan hubungan fungsional antara variabel prediktor dengan skala besar dan variabel respon dengan skala kecil. Di samping itu juga terdapat pendekatan *statistical-dynamical downscaling* yang merupakan kombinasi dari kedua metode *downscaling*.

Pendekatan *statistical downscaling* menggunakan data regional atau global untuk memperoleh hubungan fungsional antara skala lokal dengan skala global luaran GCM. Hubungan fungsional yang dimaksud adalah seperti model regresi. Pendekatan *statistical downscaling* disusun berdasarkan adanya

hubungan antara variabel prediktor dengan *grid* skala besar dan variabel respon dengan *grid* skala lokal yang dinyatakan dengan model statistik sehingga dapat digunakan untuk menerjemahkan penyimpangan-penyimpangan yang terjadi pada skala global menjadi penyimpangan dari beberapa variabel iklim lokal (Zorita dan Storch, 1999). Namun, data GCM tidak dapat digunakan secara langsung sebagai variabel prediktor pada model regresi biasa. Data luaran GCM secara umum merupakan data spasial dan temporal yang memiliki kemungkinan besar untuk terjadi korelasi spasial antar data pada *grid* yang berbeda dalam satu domain dan otokorelasi pada data deret waktu. Di samping itu domain GCM merupakan salah satu faktor dalam *statistical downscaling*, yaitu faktor lokasi dan luasan area permukaan atmosfer. Domain sangat menentukan keakuratan pendugaan model sehingga penentuan domain merupakan langkah penting dalam melakukan *statistical downscaling*. Dalam pemodelan *statistical downscaling* domain adalah sebagai variabel prediktor yang berdimensi banyak yang besar kemungkinan terjadi *curse of dimensionality*, korelasi spasial antar *grid* dalam domain dan multikolinearitas antar variabel prediktor. Semakin besar domain dan semakin banyak variabel yang digunakan akan mengakibatkan model semakin kompleks.

Pendekatan *statistical downscaling* memanfaatkan data GCM untuk melakukan peramalan iklim lokal. Hasil dari model *statistical downscaling* secara langsung berkaitan dengan statistik iklim pada waktu sebelumnya yang dapat memberikan ramalan untuk waktu yang akan datang dalam berbagai perubahan iklim. Menurut Busuioc, Chen dan Hellstrom (2001) dalam Wigena (2006), model *statistical downscaling* akan memberikan hasil yang baik apabila memenuhi beberapa syarat, antara lain:

1. Hubungan yang kuat antara variabel respon dan variabel prediktor yang menjelaskan keragaman iklim lokal dengan baik.
2. Variabel prediktor disimulasikan dengan baik oleh model GCM

3. Hubungan antara variabel respon dan variabel prediktor tidak berubah seiring dengan perubahan waktu dan tetap sama meskipun terdapat perubahan iklim.

Model *statistical downscaling* merupakan suatu fungsi yang menggambarkan hubungan fungsional sirkulasi atmosfer global dengan unsur-unsur iklim lokal. Secara umum model *statistical downscaling* adalah sebagai berikut.

$$Y_{(t,p)} = f(X_{t,q,s,g}) \quad (2.1)$$

dimana Y adalah variabel-variabel iklim lokal dan X adalah variabel-variabel dari model GCM. Adapaun t merupakan banyaknya waktu, p merupakan banyaknya variabel Y , q merupakan banyaknya variabel X , s merupakan banyaknya lapisan atmosfer dan g merupakan banyaknya *grid* domain GCM.

Secara umum model *statistical downscaling* menggunakan data deret waktu dan data spasial. Banyaknya variabel y , variabel x dan lapisan atmosfer dalam model serta adanya multikolinearitas dan otokorelasi pada variabel y maupun variabel x menunjukkan tingkat kompleksitas model. Semakin banyak variabel y dan variabel x yang digunakan, maka model *statistical downscaling* juga semakin kompleks.

2.3 *Projection Pursuit* (PP)

Projection Pursuit (PP) merupakan salah satu metode reduksi dimensi yang bersifat *underdispersed* berdasarkan pencarian suatu proyeksi informasi utama dari data berdimensi besar (Wigena, 2006). Tujuan PP adalah untuk mendapatkan informasi yang sama dari data yang berdimensi besar melalui proyeksi ke data yang berdimensi kecil dengan memaksimumkan indeks proyeksi. Reduksi dimensi menggunakan PP digunakan untuk mengatasi data yang berdimensi tinggi dan data yang tidak linier.

Prosedur reduksi dimensi menggunakan PP dilakukan berdasarkan indeks proyeksi optimum, berbeda dengan reduksi dimensi menggunakan PCA yang dilakukan berdasarkan varians

terbesar. Hasil reduksi dimensi PP kemudian digunakan untuk melakukan pemodelan *Projection Pursuit Regression* (PPR), sedangkan hasil reduksi dimensi PCA digunakan untuk melakukan pemodelan *Principal Component Regression* (PCR). Apabila $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$ adalah matriks variabel prediktor berdimensi p maka proyeksi linier $\mathcal{R}^p \rightarrow \mathcal{R}^k$ adalah sebagai berikut.

$$\mathbf{Z}^T = \mathbf{A}\mathbf{X}^T, \mathbf{X} \in \mathcal{R}^p, \mathbf{Z} \in \mathcal{R}^k, k < p \quad (2.2)$$

dimana matriks \mathbf{A} merupakan matriks proyeksi berukuran $k \times p$ dengan pangkat k . Matriks \mathbf{A} bersifat orthonormal. Apabila matriks \mathbf{X} merupakan variabel prediktor berdimensi p dengan distribusi F , maka matriks \mathbf{Z} akan berdimensi k dengan distribusi F_A .

Reduksi dimensi dengan PP menggunakan indeks proyeksi $\mathbf{I}(\mathbf{F}_A)$ untuk mendapatkan matriks \mathbf{A} yang dimaksimumkan melalui optimisasi numerik terhadap parameter. Indeks proyeksi tersebut bersifat *invariant*, yang artinya indeks proyeksi tidak bergantung pada penskalaan dan translasi.

$$\mathbf{I}(\mathbf{Z}) = \mathbf{I}(\mathbf{A}\mathbf{X}^T) = \mathbf{I}(\mathbf{A}) \quad (2.3)$$

Adapun indeks proyeksi $\mathbf{I}(\mathbf{A})$ dapat diperoleh dari persamaan berikut.

$$\mathbf{I}(\mathbf{A}) = 1 - \frac{\sum_{i=1}^t (r_i - S_\alpha(\mathbf{A}\mathbf{X}))^2}{\sum_{i=1}^t r_i^2} \quad (2.4)$$

dimana r_i merupakan residual dari pemodelan regresi OLS antara variabel respon dengan variabel prediktor. Indeks proyeksi optimum didapatkan dari selisih antara residual dan fungsi pemulus dari pemodelan.

2.4 *Projection Pursuit Regression* (PPR)

Projection Pursuit Regression (PPR) merupakan salah satu metode regresi nonparametrik dan nonlinier untuk memproses data

dalam dimensi besar yang dapat menggambarkan informasi dalam dimensi kecil melalui proses proyeksi. Dengan demikian PPR dapat mengatasi masalah rata-rata lokal, fungsi *polynomial*, dan *recursive partitioning*.

Pemodelan PPR diawali dengan memaksimumkan indeks proyeksi, menentukan fungsi-fungsi variabel tunggal secara empirik berdasarkan proyeksi-proyeksi optimum, serta menjumlahkan fungsi-fungsi tersebut (Jones dan Sibson, 1987 dalam Wigena, 2006).

Berdasarkan Friedman dan Stuetzle (1981), algoritma PPR adalah sebagai berikut.

1. Menentukan Nilai awal residual dan nilai M (banyaknya fungsi).

$$\begin{aligned} r_i &\leftarrow y_i, & i = 1, 2, \dots, t \\ M &\leftarrow 0 \end{aligned} \quad (2.5)$$

dimana $\sum y_i = 0$ (variabel respon yang distandarkan).

Adapaun proses standarisasi y_i adalah sebagai berikut.

$$z_i = \left(\frac{y_i - \mu}{\sigma} \right) \quad (2.6)$$

Setelah y_i distandarkan, langkah selanjutnya adalah meregresikan antara z_i (variabel respon) dengan \mathbf{X} (variabel prediktor) dengan menggunakan metode *Ordinary Least Square* (OLS).

$$z_i = \beta_0 + \sum_{k=1}^n \beta_k \mathbf{X}_{ik} + \varepsilon_i \quad (2.7)$$

dimana y merupakan variabel respon, \mathbf{X} merupakan variabel prediktor, i merupakan banyaknya pengamatan ($i = 1, 2, \dots, t$), k merupakan banyaknya variabel prediktor ($k = 1, 2, \dots, n$). Sedangkan μ adalah rata-rata variabel respon dan σ adalah standar deviasi variabel respon.

Residual (r_i) yang merupakan hasil regresi antara z_i dengan \mathbf{X} akan digunakan untuk menentukan indeks proyeksi minimum dalam model PPR. Sedangkan penentuan banyaknya fungsi didapatkan dari *scatterplot* antara variabel respon dan variabel prediktor. Banyaknya fungsi ditentukan berdasarkan optimalisasi banyak fungsi $m=1,2,3,4,5$. Banyak fungsi yang dipilih ditentukan berdasarkan hasil validasi terbaik.

2. Menentukan α dan S_α dalam model.

Untuk kombinasi linier $\mathbf{Z} = \alpha_m \mathbf{X}$, ditentukan fungsi pemulus $S_\alpha(\mathbf{Z})$ sesuai dengan nilai-nilai \mathbf{Z} dengan menggunakan indeks proyeksi $\mathbf{I}(\alpha)$. Adapun indeks proyeksi $\mathbf{I}(\alpha)$ dapat dihitung dengan menggunakan rumus berikut.

$$\mathbf{I}(\alpha) = 1 - \frac{\sum_{i=1}^t (r_i - S_\alpha(\alpha x_i))^2}{\sum_{i=1}^t r_i^2} \quad (2.8)$$

Menentukan vektor koefisien α_{M+1} yang memaksimumkan $\mathbf{I}(\alpha)$ (*projection pursuit*), $\alpha_{M+1} = \max_{\alpha}^{-1}(\mathbf{I}(\alpha))$ dan fungsi pemulusnya adalah $S_{\alpha_{M+1}}(\mathbf{Z})$.

3. Akhir algoritma.

Jika $\mathbf{I}(\alpha)$ lebih kecil dari nilai *threshold*, maka algoritma PPR berhenti. Namun, jika $\mathbf{I}(\alpha)$ lebih besar dari nilai *threshold*, maka nilai residual dan nilai M dirubah.

$$\begin{aligned} r_i &\leftarrow r_i - S_\alpha(\mathbf{Z}), & i = 1, 2, \dots, t \\ M &\leftarrow M + 1 \end{aligned} \quad (2.9)$$

Kemudian kembali ke langkah 2. Nilai *threshold* didapatkan berdasarkan batas antara kombinasi linier dalam *scatterplot* antara variabel respon dan variabel prediktor.

Fungsi pemulus $S_\alpha(\mathbf{Z})$ ditentukan secara nonparametrik. Bentuk umum hubungan antara variabel respon dan \mathbf{Z} dengan fungsi pemulusnya secara matematis dapat dituliskan sebagai berikut.

$$y_i = S_\alpha(z_i) + r_i \quad (2.10)$$

Model regresi secara umum dapat ditulis sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i \quad (2.11)$$

dimana ε_i adalah IID dengan $E(\varepsilon_i) = 0$ dan $f(x)$ kontinyu. Dalam regresi nonparametrik fungsi $f(x)$ diduga dengan $S_\alpha(z)$, yang ditentukan berdasarkan rata-rata lokal, yaitu:

$$S(y_i) = \text{AVE}_{i-k \leq j \leq i+k}(y_j) \quad (2.12)$$

untuk lebar jendela (*bandwidth*) k tertentu dan dengan formulasi AVE seperti median atau rata-rata. Pemilihan nilai k sangat menentukan varians penduga dan besarnya bias. Nilai k yang terlalu kecil akan memperbesar varians penduga, sedangkan nilai k yang terlalu besar akan memperbesar bias.

Adapun penentuan fungsi pemulus $S_\alpha(z)$ menurut Friedman dan Stuetzle (1981) adalah sebagai berikut.

1. Menentukan median untuk setiap tiga respon secara sekuensial untuk menghilangkan pengaruh data pencilan.
2. Menentukan penduga varians respon pada setiap titik dengan residual kuadrat rata-rata (*average squared residual*) dari penduga linier lokal ke k tertentu.
3. Pemulusan penduga varians dengan rata-rata bergerak dan k tetap untuk menghindari perhitungan lebih dari satu kombinasi linier $\mathbf{Z} = \boldsymbol{\alpha}_m \mathbf{X}$.
4. Pemulusan berturut dari tahap 1 dengan pengepasan (*fitting*) linier lokal dengan nilai k yang diperoleh pada tahap 3.

Model PPR secara matematis diuraikan berdasarkan fungsi kernel (*kernel-based PPR*) dan sifat penduga PP. Pada dasarnya solusi PPR *invariant* terhadap setiap transformasi, baik dalam

rotasi maupun dalam penskalaan variabel prediktor. Apabila $S(z)$ merupakan pemetaan $\mathfrak{R}^p \rightarrow \mathfrak{R}$, $f(x)$ merupakan fungsi kepekatan dalam \mathfrak{R}^p , dan \mathbf{X} adalah variabel acak berdimensi p , maka penduga PP yang digunakan untuk mendapatkan proyeksi utama untuk suatu skalar z adalah sebagai berikut.

$$S_\alpha(z) = E\{f(x) | \alpha\mathbf{X} = z\} \quad (2.13)$$

Proyeksi pertama terhadap $f(x)$ adalah fungsi $f_1(x) = S_{\alpha_1}(z)$, dimana α_1 akan meminimumkan persamaan $L(\alpha)$ berikut.

$$L(\alpha) = E\left[\{f(x) - S_\alpha(z)\}^2\right] \quad (2.14)$$

sehingga persamaan $L(\alpha)$ menjadi

$$\hat{L}(\alpha) = \frac{1}{t} \sum_{i=1}^t \{y_k - S_{\alpha_i}(\alpha_i \mathbf{X})\}^2 \quad (2.15)$$

dengan penduga proyeksi pertamanya adalah

$$\hat{f}_1(x) = \hat{S}_{\alpha_1}(\hat{\alpha}_1 \mathbf{X}) \quad (2.16)$$

Proses proyeksi ke- m memiliki tahapan yang sama seperti langkah-langkah mendapatkan proyeksi ke-1. Penduga $\hat{S}_\alpha(z)$ akan konsisten dan konvergen terhadap $S_\alpha(z)$, $\hat{\alpha}$ juga akan konvergen terhadap α , sehingga didapatkan model PPR sebagai berikut.

$$y_i = \sum_{m=1}^M S_{\alpha_m}(\alpha_m \mathbf{X}) = \bar{y} + \sum_{m=1}^M \beta_m f_m \left(\sum_{k=1}^p \alpha_{km} X_k \right) + \varepsilon_i \quad (2.17)$$

dimana:

- y_i = variabel respon
- \bar{y} = rata-rata variabel respon
- $S_{\alpha_m}(\alpha_m \mathbf{X})$ = suatu fungsi yang tidak diketahui
- α_m = vektor satuan (arah *projection pursuit*)

	$= (\alpha_{1m}, \alpha_{2m}, \dots, \alpha_{km})$
\mathbf{X}	$= \text{variabel prediktor} = (X_{i1}, X_{i2}, \dots, X_{ik})$
ε_i	$= \text{faktor acak dengan } E(\varepsilon_i) = 0 \text{ dan } \text{var}(\varepsilon_i) = \sigma^2$
m	$= \text{banyak fungsi}$
i	$= \text{satuan waktu}$
k	$= \text{banyak variabel prediktor}$
X_i dan ε_i bebas.	

2.5 Validasi Model

Validasi model merupakan suatu proses untuk mengetahui apakah model yang terbentuk menghasilkan nilai prediksi yang akurat. Dalam melakukan pemodelan data dibagi menjadi dua bagian, yaitu data *in-sample* (data *training*) dan data *out-sample* (data *testing*). Data *in-sample* digunakan untuk mendapatkan model, sedangkan data *out-sample* digunakan untuk melakukan validasi terhadap model tersebut. Pemilihan model terbaik berdasarkan data *in-sample* dilakukan menggunakan RMSE, sedangkan data *out-sample* dilakukan menggunakan RMSEP. Selain itu pemilihan model terbaik berdasarkan data *out-sample* juga dapat dilakukan menggunakan $R^2_{\text{prediction}}$ atau koefisien determinasi. Semakin kecil nilai RMSE, RMSEP menunjukkan bahwa model yang didapatkan semakin baik. Sebaliknya, semakin besar nilai $R^2_{\text{prediction}}$ menunjukkan bahwa model yang didapatkan semakin baik.

RMSE (*Root Mean Square Error*) merupakan akar kuadrat dari rata-rata kuadrat *error*, dimana *error* adalah selisih dari data *in-sample* dan data hasil pemodelan. Kriteria RMSE yaitu membandingkan data *in-sample* dengan taksiran data hasil pemodelan. Adapaun RMSE secara matematis dapat dituliskan sebagai berikut.

$$RMSE = \sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} (Y_i - \hat{Y}_i)^2} \quad (2.18)$$

dimana Y_i merupakan data *in-sample* ke- i , \hat{Y}_i merupakan taksiran data hasil pemodelan data *in-sample* ke- i dan n_1 meruakan jumlah data *in-sample*.

RMSEP (*Root Mean Square Error Prediction*) merupakan akar kuadrat dari rata-rata kuadrat *error*, dimana *error* adalah selisih dari data *out-sample* dan data hasil peramalan. Sedangkan $R^2_{prediction}$ merupakan proporsi keragaman total nilai-nilai variabel respon yang dapat diterangkan oleh variabel prediktor dalam model yang digunakan. Kriteria RMSEP dan $R^2_{prediction}$ dilakukan dengan membandingkan data *out-sample* dengan data hasil peramalan. Adapun RMSEP dan $R^2_{prediction}$ secara berturut-turut dapat dihitung menggunakan rumus sebagai berikut.

$$RMSEP = \sqrt{\frac{1}{n_2} \sum_{j=1}^{n_2} (Y_j - \hat{Y}_j)^2} \quad (2.19)$$

$$R^2_{prediction} = \left(\frac{\sum_{j=1}^{n_2} (\hat{Y}_j - \bar{Y})^2}{\sum_{j=1}^{n_2} (Y_j - \bar{Y})^2} \right) \times 100\% \quad (2.20)$$

dimana Y_j merupakan data *out-sample* ke- j , \hat{Y}_j merupakan taksiran data hasil pemodelan data *out-sample* ke- j , \bar{Y} merupakan rata-rata variabel respon dan n_2 meruakan jumlah data *out-sample*.

2.6 Curah Hujan

Presipitasi merupakan hasil dari kondensasi uap air di atmosfer yang terjadi ketika atmosfer menjadi jenuh dan air terkondensasi. Presipitasi dapat berwujud padat (salju dan hujan es), aerosol (embun dan kabut) atau cairan (hujan). Hujan merupakan salah satu bentuk presipitasi yang berwujud cairan yang dapat

mencapai permukaan bumi (Sosrodarsono dan Takeda, 2003). Hujan terbentuk apabila titik air terpisah jatuh dari awan ke bumi. Curah hujan didefinisikan sebagai banyaknya air yang jatuh ke permukaan bumi. Intensitas curah hujan diartikan sebagai jumlah hujan dalam satu satuan waktu yang dinyatakan dalam satuan milimeter per jam (mm/jam) (LAPAN, 2014a). Menurut Sosrodarsono dan Takeda (2003), hujan dapat digolongkan menjadi beberapa golongan berdasarkan intensitas curah hujannya. Penggolongan hujan berdasarkan intensitas curah hujannya dapat dilihat pada Tabel 2.1.

Tabel 2.1 Derajat Hujan Berdasarkan Intensitas Curah Hujannya

Derajat Hujan	Intensitas Curah Hujan (mm/menit)
Hujan sangat lemah	$<0,02$
Hujan lemah	$0,02 < CH \leq 0,05$
Hujan normal	$0,05 < CH \leq 0,25$
Hujan deras	$0,25 < CH \leq 1,00$
Hujan sangat deras	$>1,00$

Secara umum curah hujan di Indonesia dipengaruhi oleh beberapa fenomena alam, antara lain gelombang El-Nino, monsun Asia-Australia, sirkulasi atmosfer lokal, regional dan global, seperti sirkulasi utara-selatan, sirkulasi barat-timur dan sistem angin lokal. Apabila salah satu dari sistem sirkulasi tersebut terganggu, maka secara otomatis akan mempengaruhi cuaca dan iklim di Indonesia (LAPAN, 2014a). Curah hujan memiliki keragaman yang tinggi dalam ruang dan waktu. Keragaman curah hujan berdasarkan skala ruang dipengaruhi oleh dua faktor, yaitu ketinggian tempat dan angin. Semakin rendah ketinggian suatu tempat maka potensi curah hujan yang diterima akan semakin banyak dikarenakan tempat tersebut memiliki suhu yang rendah. Adapaun faktor angin dibedakan menjadi dua, yaitu angin *zonal* dan angin *meridional*. Angin *zonal* merupakan komponen angin barat-timur yang berpengaruh terhadap fenomena *Dipole Mode* (DM) yang terjadi dikarenakan adanya pergerakan massa udara dari barat ke timur Samudera Hindia ataupun sebaliknya. Angin

meridional berpengaruh terhadap fenomena DM sebagai akibat dari adanya aliran udara antara wilayah India bagian selatan dengan wilayah Australia bagian selatan barat. Adapun *Dipole Mole* (DM) merupakan gejala naiknya suhu permukaan laut dari kondisi normal di sepanjang ekuator Samudera Hindia khususnya di sebelah selatan India yang diikuti dengan menurunnya suhu permukaan laut pada kondisi tidak normal di perairan Indonesia khususnya di wilayah barat Pulau Sumatera (LAPAN, 2014b).

Kondisi El-Nino dan La-Nina terjadi berdasarkan pada nilai anomali suhu permukaan laut di daerah NINO di Samudera Pasifik dekat ekuator bagian tengah dan timur. El-Nino merupakan fenomena global dari sistem interaksi laut dan atmosfer yang ditandai dengan memanasnya suhu permukaan air laut di Pasifik ekuator, dimana suhu permukaan air laut di daerah tersebut lebih panas dari suhu rata-ratanya. Fenomena alam El-Nino secara umum akan menyebabkan curah hujan berkurang di sebagian besar wilayah Indonesia, berkurangnya curah hujan bergantung pada lokasi dan intensitas El-Nino tersebut. Namun, tidak seluruh wilayah Indonesia terkena dampak fenomena El-Nino dikarenakan luasnya wilayah Indonesia dan posisi topografisnya yang dikenal sebagai benua maritim. Sebaliknya, fenomena La-Nina ditandai dengan naiknya suhu permukaan laut di Samudera Pasifik tengah dan timur, dimana suhu permukaan air laut di daerah tersebut lebih panas dari suhu rata-ratanya, sehingga mengakibatkan curah hujan bertambah di wilayah Indonesia.

2.7 Penelitian Sebelumnya

Penelitian yang dilakukan menggunakan metode *statistical downscaling* telah banyak dilakukan dalam bidang meteorologi di Indonesia, antara lain Wigena (2006) yang melakukan penelitian mengenai prediksi curah hujan bulanan di Indramayu menggunakan pemodelan *statistical downscaling* dengan pendekatan *Projection Pursuit Regression* (PPR). Selain itu Anitawati (2010) juga pernah melakukan penelitian mengenai pemodelan *statistical downscaling* dengan metode *Principal Component Regression* (PCR) dan *Projection Pursuit regression*

(PPR), dimana hasil penelitian tersebut menunjukkan bahwa pemodelan dengan metode PPR memiliki hasil yang lebih baik dibandingkan dengan metode PCR. Metode PPR tidak hanya dilakukan untuk melakukan peramalan cuaca jangka panjang tetapi juga digunakan untuk meramalkan cuaca jangka pendek. Salah satunya adalah penelitian yang dilakukan oleh Safitri (2012) yaitu meramalkan suhu minimum, suhu maksimum dan kelembaban menggunakan *Model Output Statistics* (MOS) dengan metode PPR. Berdasarkan berbagai penelitian tersebut metode PPR menghasilkan model yang baik karena PPR dapat mengatasi berbagai permasalahan dalam melakukan peramalan cuaca, seperti multikolinearitas antar variabel prediktor dan sifat variabel respon yang nonlinier serta tidak berdistribusi baku.

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Meteorologi, Klimatologi dan Geofisika (BMKG) periode Januari 1991 sampai Desember 2010. Data tersebut merupakan data observasi unsur iklim yaitu berupa data curah hujan bulanan. Adapun data GCM diperoleh dari website resmi *Royal Netherlands Meteorological Institute* (KNMI), yaitu <http://climexp.knmi.nl>. Wilayah yang menjadi lokasi penelitian adalah lima kabupaten yang menjadi sentra produksi padi Jawa Timur dengan masing-masing empat stasiun pengamatan seperti yang disajikan pada Tabel 3.1 berikut.

Tabel 3.1 Lokasi Stasiun Pengamatan

No	Kabupaten	Nama Stasiun Pengamatan
1	Ngawi	Karangjati, Walikukun, Mantingan, Ngale
2	Lamongan	Lamongan, Babat, Mantup, Brondong
3	Bojonegoro	Jatiblimbing, Kedungadem, Dander, Klepek
4	Jember	Ledokombo, Pakusari, Sukowono, Ajung
5	Banyuwangi	Kawah Ijen, Kabat, Pasewaran, Bajulmati

Stasiun pengamatan di atas dipilih karena pada stasiun tersebut pengukuran cuaca dilakukan secara rutin sehingga data pengamatan yang dihasilkan relatif baik dan lengkap dibandingkan dengan stasiun lainnya.

3.2 Variabel Penelitian

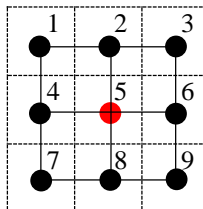
Variabel respon yang digunakan dalam penelitian ini adalah curah hujan dalam satuan waktu bulanan. Curah hujan tersebut diukur secara langsung di empat stasiun pada lima kabupaten yang menjadi sentra produksi padi Jawa Timur. Kemudian data hasil pengukuran curah hujan di empat stasiun pengamatan tersebut dihitung nilai rata-ratanya. Rata-rata curah hujan tersebut yang kemudian digunakan sebagai variabel respon.

Variabel prediktor yang digunakan adalah luaran GCM yaitu aplikasi model NCEP CFSR (*National Centers for Environmental Prediction - Climate Forecast System Reanalysis*) dengan resolusi sebesar $0,3^0 \times 0,3^0$ atau kurang lebih setara dengan $33 \text{ km} \times 33 \text{ km}$. Variabel GCM dari aplikasi model NCEP CFSR yang digunakan sebanyak 4 variabel. Penjelasan mengenai variabel prediktor dapat dilihat pada Tabel 3.2.

Tabel 3.2 Variabel GCM dari Aplikasi Model NCEP CFSR

No	Variabel	Keterangan	Satuan
1	<i>Temperature</i>	Suhu	K
2	<i>Precipitation</i>	Presipitasi (hujan)	mm
3	<i>Zonal Wind</i>	Angin <i>zonal</i>	ms^{-1}
4	<i>Meridional Wind</i>	Angin <i>meridional</i>	ms^{-1}

Semua variabel tersebut diukur pada *grid* 3×3 . Adapun sistem pemberian nomor *grid* dalam NCEP ESFR dapat dilihat pada Gambar 3.1.



Gambar 3.1 Penentuan Pemilihan *Grid* di Setiap Lokasi Pengamatan (BMKG, 2011)

Bagian yang berwarna merah (*grid* nomor 5) adalah *grid* NCEP CFSR yang tepat berada di kabupaten yang menjadi lokasi pengamatan. Dalam variabel luaran GCM, setiap *grid* merupakan variabel prediktor sehingga pada domain *grid* 3×3 terdapat 9 variabel prediktor. Oleh karena itu, jumlah variabel prediktor yang digunakan dalam pemodelan adalah sebanyak 36 variabel prediktor.

3.3 Struktur Data

Struktur data yang digunakan dalam penelitian ini disajikan pada Tabel 3.3.

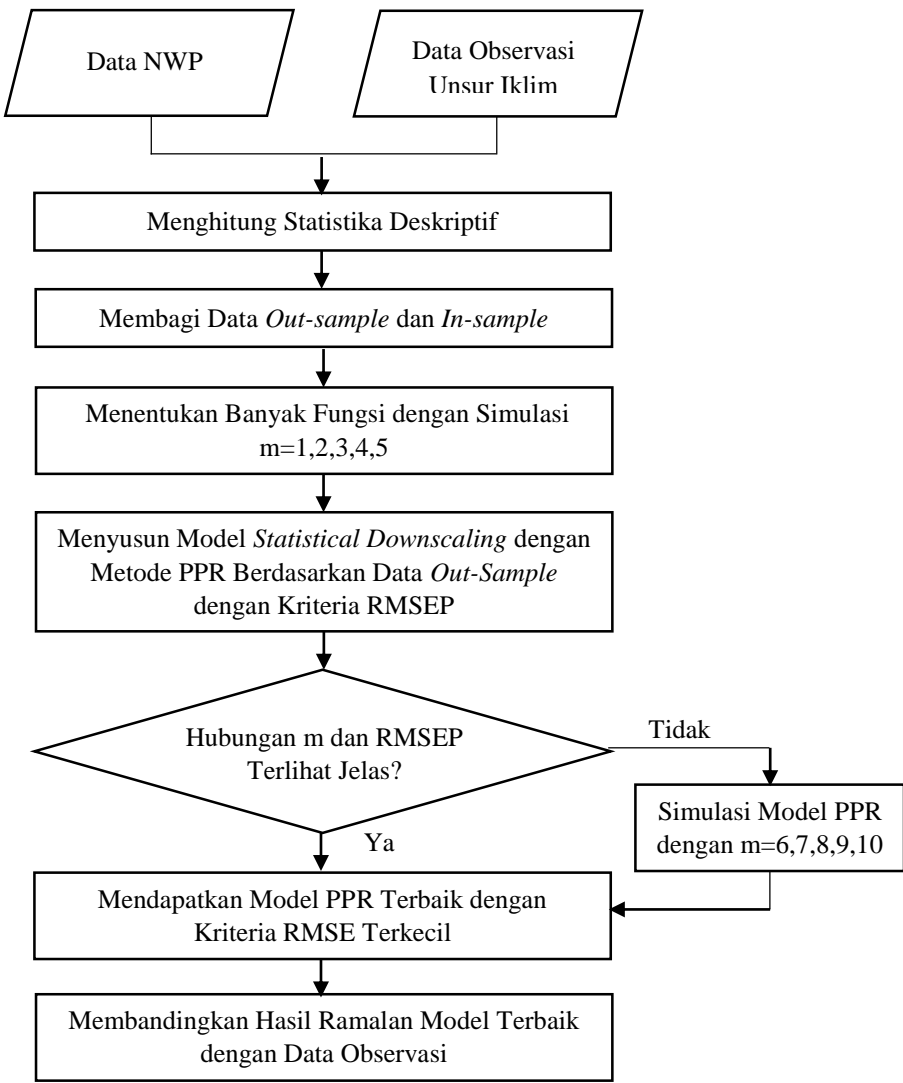
Tabel 3.3 Struktur Data Penelitian

No	Respon	Prediktor						
	Curah Hujan	<i>Temperature</i>			...	<i>Meridional Wind</i>		
		1	...	9	...	1	...	9
1	y ₁	X _{1,1,1}	...	X _{1,9,1}	...	X _{4,1,1}	...	X _{4,9,1}
2	y ₂	X _{1,1,2}	...	X _{1,9,2}	...	X _{4,1,2}	...	X _{4,9,2}
3	y ₃	X _{1,1,3}	...	X _{1,9,3}	...	X _{4,1,3}	...	X _{4,9,3}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	y _n	X _{1,1,n}	...	X _{1,9,n}	...	X _{4,1,n}	...	X _{4,9,n}

3.4 Langkah Analisis Data

Langkah analisis yang digunakan dalam melakukan penelitian ini adalah sebagai berikut.

1. Mendeskripsikan data.
 - a. Membuat *boxplot* untuk mengetahui apakah terdapat data *outlier* serta mengetahui rata-rata tertinggi dan terendah di masing-masing lokasi penelitian.
 - b. Membuat *time series plot* untuk mengidentifikasi pola variabel respon.
 - c. Membuat *scatterplot* untuk mengetahui hubungan antara variabel respon dan variabel prediktor.
2. Membentuk model *statistical downscaling* menggunakan metode *Projection Pursuit Regression* (PPR).
 - a. Membagi data menjadi data *in-sample* dan *out-sample*. Pembagian data dilakukan dengan pembagian proporsi sebesar 90:10, artinya data *in-sample* sebanyak 90% dan data *out-sample* sebanyak 10%.
 - b. Menentukan banyaknya fungsi dalam model PPR. Banyaknya fungsi ditentukan dari simulasi $m=1,2,3,4,5$.
 - c. Membuat model PPR berdasarkan data *in-sample* dengan simulasi banyak fungsi $m=1,2,3,4,5$.



Gambar 3.2 Diagram Alir Pemodelan *Statistical Downscaling* dengan PPR

3. Melakukan validasi terhadap model *statistical downscaling*.
Validasi model dilakukan untuk mengetahui keakuratan model *statistical downscaling* berdasarkan kriteria nilai RMSEP. Validasi model berdasarkan data *out-sample* dilakukan pada setiap model *statistical downscaling* dengan banyak fungsi $m=1,2,3,4,5$.
4. Mengidentifikasi hubungan m dengan RMSEP.
Identifikasi dilakukan untuk mencari banyak fungsi (m) yang memiliki RMSEP terkecil. Jika pola hubungan antara banyak fungsi (m) dengan RMSEP sudah terlihat jelas (misalnya banyak fungsi berbanding lurus atau terbalik dengan nilai RMSEP), maka simulasi fungsi hanya dilakukan sampai $m=5$. Namun, apabila pola hubungan belum terlihat dengan jelas maka simulasi dilanjutkan sampai $m=10$. Semakin banyak fungsi (m) yang digunakan tidak bisa menjamin bahwa model yang didapatkan akan semakin baik. Sebaliknya semakin banyak fungsi (m) yang digunakan, model yang didapatkan juga akan semakin kompleks.
5. Mendapatkan model *statistical downscaling* terbaik.
Model *statistical downscaling* terbaik adalah model dengan banyak fungsi tertentu yang memiliki nilai RMSEP terkecil.
Tahapan analisis data selengkapnya disajikan dalam diagram alir pada Gambar 3.2.

(Halaman ini sengaja dikosongkan)

BAB IV

ANALISIS DAN PEMBAHASAN

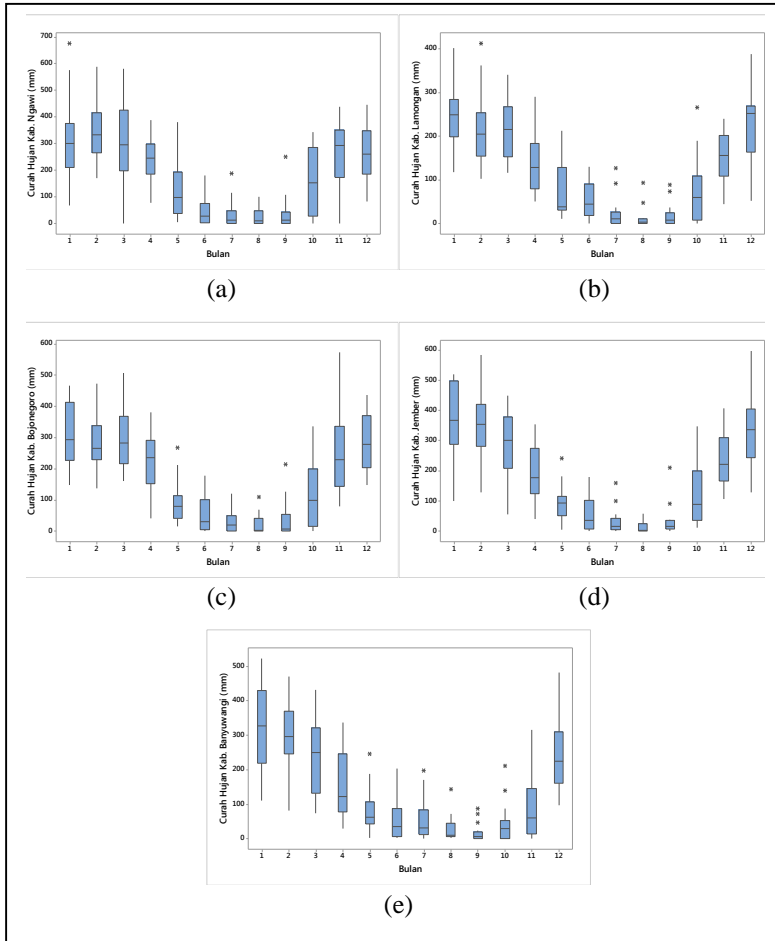
Pada bab ini membahas penyusunan model *statistical downscaling* dengan metode *Prejection Pursuit Regression* (PPR) untuk meramalkan curah hujan bulanan di sentra produksi padi Jawa Timur, yaitu Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi. Sebanyak 90 persen data digunakan untuk membuat model, sedangkan 10 persen data digunakan untuk validasi model. Bagian awal dilakukan deskripsi curah hujan di lima kabupaten dan bagian akhir dilakukan pemilihan model terbaik untuk masing-masing Kabupaten.

4.1 Deskripsi Curah Hujan di Sentra Produksi Padi Jawa Timur

Curah hujan di lima kabupaten yang menjadi sentra produksi padi Jawa Timur memiliki karakteristik yang hampir sama. Berdasarkan *boxplot* curah hujan pada Gambar 4.1 menunjukkan bahwa terdapat data *outlier* pada seluruh wilayah penelitian. Bulan Januari merupakan bulan yang memiliki rata-rata curah hujan tertinggi untuk Kabupaten Lamongan, Bojonegoro, Jember dan Banyuwangi, sedangkan untuk Kabupaten Ngawi rata-rata curah hujan tertinggi terjadi pada bulan Februari. Hal tersebut sesuai dengan tipe hujan di wilayah Jawa Timur yang memiliki satu puncak musim hujan di mana biasanya terjadi pada bulan Desember, Januari atau Februari. Rata-rata curah hujan tertinggi di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi secara berturut-turut adalah 342,3 mm; 250,3 mm; 312,0 mm; 372,9 mm dan 323,8 mm.

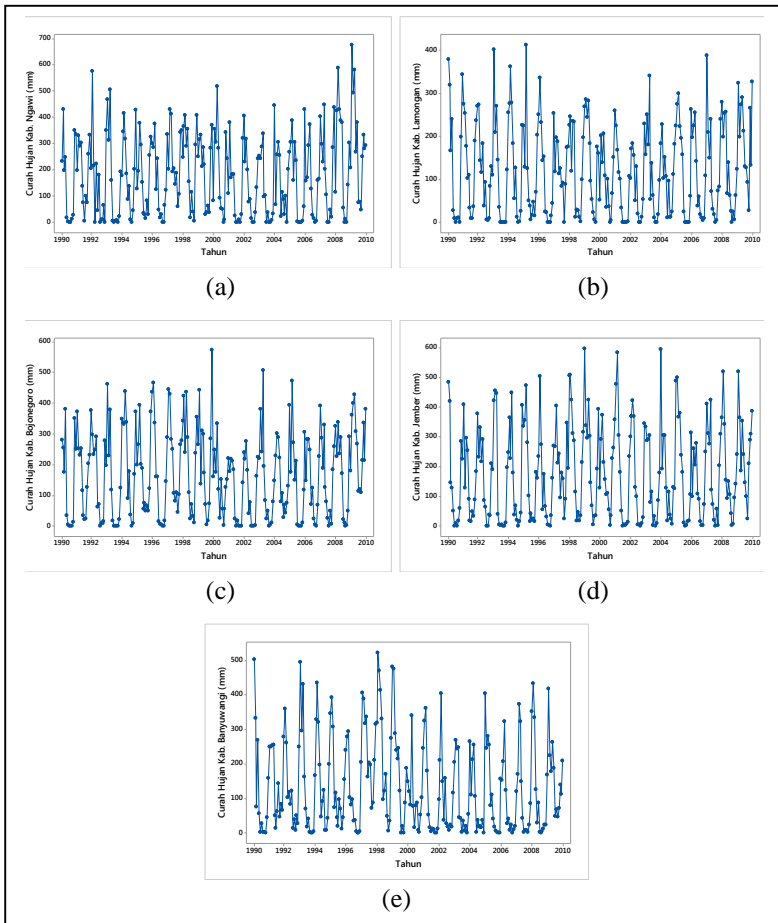
Rata-rata curah hujan terendah di Kabupaten Ngawi, Lamongan, Bojonegoro dan Jember terjadi pada bulan Agustus, sedangkan di Kabupaten Banyuwangi rata-rata curah hujan terendah terjadi pada bulan September. Hal tersebut sesuai dengan iklim di wilayah Jawa Timur yang memiliki satu puncak musim kemarau di mana biasanya terjadi pada bulan Agustus atau

September. Rata-rata curah hujan terendah secara berturut-turut di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi adalah 25,25 mm; 10,41 mm; 21,07 mm; 13,14 mm dan 15,48 mm.



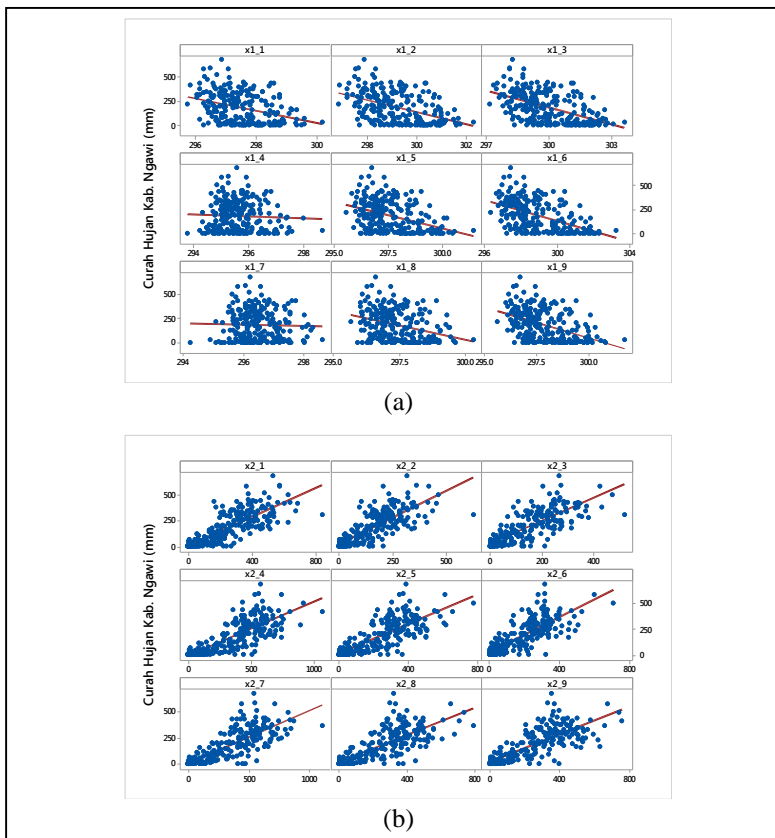
Gambar 4.1 *Boxplot* Curah Hujan di Wilayah Pengamatan (a) Kabupaten Ngawi, (b) Kabupaten Lamongan, (c) Kabupaten Bojonegoro, (d) Kabupaten Jember dan (e) Kabupaten Banyuwangi

Berdasarkan *time series plot* curah hujan pada Gambar 4.2 dapat dilihat bahwa curah hujan di lima kabupaten sentra produksi padi Jawa Timur bersifat fluktuatif. Standar deviasi yang besar menunjukkan bahwa curah hujan di lima kabupaten mengalami kenaikan dan penurunan yang signifikan.

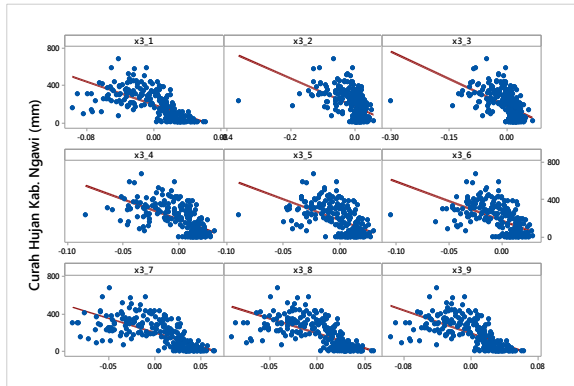


Gambar 4.2 *Time Series Plot* Curah Hujan di Wilayah Pengamatan (a) Kabupaten Ngawi, (b) Kabupaten Lamongan, (c) Kabupaten Bojonegoro, (d) Kabupaten Jember dan (e) Kabupaten Banyuwangi

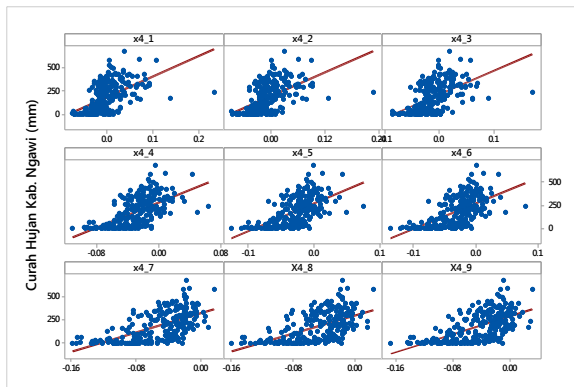
Setelah melakukan identifikasi data curah hujan (variabel respon) menggunakan *time series plot*, langkah selanjutnya yaitu melakukan identifikasi pola hubungan antara variabel respon berupa data observasi curah hujan dan variabel prediktor berupa data luaran GCM. Identifikasi pola hubungan antara variabel respon dan variabel prediktor dapat dilakukan dengan membuat *scatterplot*. Pola hubungan tersebut secara visual dapat dilihat pada Gambar 4.3.



Gambar 4.3 *Scatterplot* Antara Variabel Respon dan Variabel Prediktor di Kabupaten Ngawi (a) *Temperature* (b) *Precipitation*, (c) *Zonal Wind* dan (d) *Meridional Wind*



(c)



(d)

Gambar 4.3 Scatterplot Antara Variabel Respon dan Variabel Prediktor di Kabupaten Ngawi (a) *Temperature*, (b) *Precipitation*, (c) *Zonal Wind* dan (d) *Meridional Wind* (lanjutan)

Berdasarkan Gambar 4.3 dapat diketahui bahwa tidak terdapat pola hubungan yang linier antara variabel respon dan variabel prediktor di Kabupaten Ngawi. Variabel respon yang dimaksud yaitu berupa data observasi curah hujan dan variabel prediktor berupa data luaran GCM. Tidak adanya pola hubungan tersebut sesuai dengan karakteristik data curah hujan yang tidak linier. Oleh karena itu, pembentukan model curah hujan sebaiknya

dilakukan dengan menggunakan metode analisis non linier sehingga mendapatkan hasil ramalan yang baik. Adapun *scatterplot* untuk Kabupaten Lamongan, Bojonegoro, Jember dan Banyuwangi secara lengkap disajikan pada Lampiran B.

4.2 Pemodelan *Statistical Downscaling* dengan *Projection Pursuit Regression* di Sentra Produksi Padi Jawa Timur

Pemodelan *statistical downscaling* dengan *Projection Pursuit Regression* (PPR) dilakukan dengan meregresikan data observasi curah hujan sebagai variabel respon dan data luaran aplikasi GCM sebagai variabel prediktor. Data curah hujan yang digunakan adalah rata-rata curah hujan dari empat stasiun masing-masing di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi.

Pemodelan *statistical downscaling* dengan PPR untuk meramalkan curah hujan bulanan memiliki beberapa tahapan. Tahap pertama yaitu menentukan banyak fungsi dalam model PPR dengan optimisasi simulasi banyak fungsi $m = 1, 2, 3, 4$ dan 5 . Tahap kedua yaitu membangun model PPR antara data observasi curah hujan dan data luaran GCM. Tahap ketiga yaitu memvalidasi model PPR menggunakan data *out-sample*. Validasi model dilakukan dengan membandingkan nilai RMSEP dan R^2_{pred} pada masing-masing fungsi. Model terbaik dipilih berdasarkan nilai RMSEP terkecil dan atau nilai R^2_{pred} terbesar. Perbandingan nilai RMSEP dan R^2_{pred} di lima kabupaten yang menjadi sentra produksi padi di Jawa Timur berdasarkan banyaknya fungsi disajikan pada Tabel 4.1.

Tabel 4.1 menunjukkan bahwa nilai RMSEP tidak berpengaruh terhadap banyaknya fungsi (m). Hal tersebut memiliki arti bahwa semakin banyak fungsi yang digunakan, nilai RMSEP tidak selalu semakin besar dan nilai R^2_{pred} tidak selalu semakin kecil. Nilai RMSE terkecil di lima kabupaten didapatkan ketika banyak fungsi (m) yang digunakan adalah 1. Oleh karena itu, model optimal untuk meramalkan curah hujan di lima kabupaten didapatkan ketika $m = 1$. Nilai RMSEP yang didapatkan dengan banyak fungsi $m = 1$ di Kabupaten Ngawi dan Lamongan

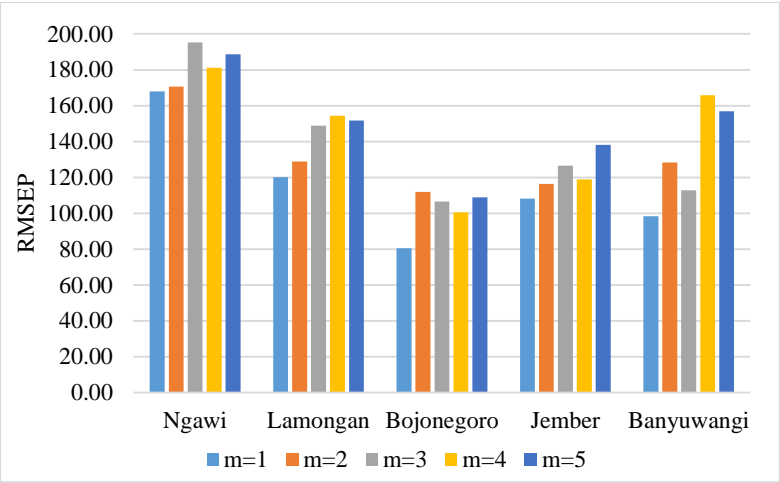
adalah sebesar 168,00 dan 120,08 dengan nilai R^2_{pred} berturut-turut sebesar 56,55 dan 43,89. RMSEP yang diperoleh ketika $m = 1$ untuk Kabupaten Bojonegoro adalah sebesar 80,44 dengan R^2_{pred} sebesar 81,23. Sedangkan RMSEP dengan banyak fungsi $m = 1$ di Kabupaten Jember dan Banyuwangi adalah sebesar 108,20 dan 98,41 dengan R^2_{pred} sebesar 74,55 untuk Kabupaten Jember dan 55,77 untuk Kabupaten Banyuwangi. Tingginya nilai RMSEP ini kemungkinan dipengaruhi oleh keragaman data curah hujan yang cukup tinggi. Di samping itu, tingginya RMSEP kemungkinan juga bisa dikarenakan adanya faktor-faktor yang mempengaruhi curah hujan yang belum tertangkap secara keseluruhan oleh variabel yang digunakan dalam penelitian ini.

Tabel 4.1 Nilai RMSEP dan R^2_{pred} Menurut Banyaknya Fungsi di Lima Kabupaten

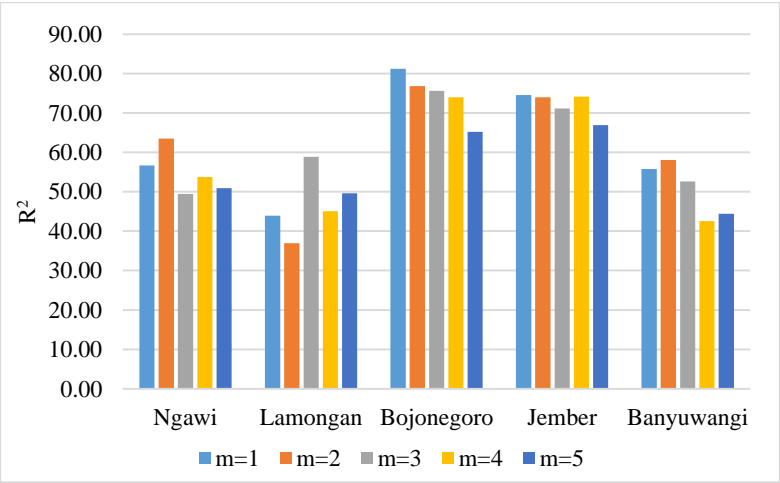
Kabupaten	Validasi Model	m=1	m=2	m=3	m=4	m=5
Ngawi	RMSEP	168,00	170,64	195,23	181,13	188,73
	R^2_{pred}	56,66	63,48	49,48	53,72	50,93
Lamongan	RMSEP	120,08	128,79	148,91	154,41	151,66
	R^2_{pred}	43,89	36,95	58,86	45,04	49,61
Bojonegoro	RMSEP	80,44	111,83	106,61	100,50	108,82
	R^2_{pred}	81,23	76,82	75,60	73,96	65,22
Jember	RMSEP	108,20	116,31	126,49	118,84	138,17
	R^2_{pred}	74,55	73,96	71,10	74,15	66,88
Banyuwangi	RMSEP	98,41	128,25	112,78	165,73	156,97
	R^2_{pred}	55,77	58,03	52,65	42,51	44,39

■ = Nilai RMSEP terkecil

Secara visual perbandingan nilai RMSEP di lima kabupaten yang menjadi sentra produksi padi di Jawa Timur berdasarkan banyaknya fungsi disajikan pada Gambar 4.4. Sedangkan perbandingan nilai R^2_{pred} berdasarkan banyaknya fungsi secara visual disajikan pada Gambar 4.5.



Gambar 4.4 Nilai RMSEP Menurut Banyaknya Fungsi di Lima Kabupaten



Gambar 4.5 Nilai R^2_{pred} Menurut Banyaknya Fungsi di Lima Kabupaten

Model *statistical downscaling* yang optimal untuk meramalkan curah hujan di Kabupaten Ngawi dengan $m = 1$ adalah sebagai berikut.

$$\hat{y} = 278,93 + 127,03 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

Matriks koefisien proyeksi α_{k1} yang memproyeksikan X_k pada persamaan model PPR yang optimal di Kabupaten Ngawi dengan $m = 1$ adalah sebagai berikut.

$$\alpha_{k1}^T = \begin{bmatrix} -0,03 \\ 0,13 \\ -0,19 \\ -0,10 \\ -0,02 \\ 0,13 \\ 0,33 \\ -0,43 \\ 0,22 \\ \vdots \\ \vdots \\ 0,13 \\ 0,55 \\ 0,12 \end{bmatrix}$$

Matriks α_{k1} berukuran 1×36 , sesuai dengan banyak fungsi dalam model yaitu $m = 1$ dan jumlah variabel prediktor sebanyak 36. Pada model optimal di Kabupaten Ngawi matriks koefisien proyeksi α_{k1} memproyeksikan X_k (variabel GCM) sepanjang 127,03. Matriks koefisien proyeksi α_{k1} secara lengkap dapat dilihat pada Lampiran C.

Model *statistical downscaling* yang optimal untuk meramalkan curah hujan di Kabupaten Lamongan dengan $m = 1$ adalah sebagai berikut.

$$\hat{y} = 151,08 + 94,36 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

Berikut merupakan matriks koefisien proyeksi α_{k1} yang memproyeksikan X_k pada persamaan model PPR optimal di Kabupaten Lamongan dengan banyak fungsi $m = 1$.

$$\alpha_{k1}^T = \begin{bmatrix} -0,15 \\ 0,00 \\ 0,23 \\ 0,18 \\ -0,30 \\ 0,03 \\ -0,02 \\ -0,01 \\ 0,02 \\ \vdots \\ \vdots \\ -0,38 \\ 0,02 \\ 0,36 \end{bmatrix}$$

Matriks α_{k1} berukuran 1×36 , sesuai dengan banyak fungsi dalam model yaitu $m = 1$ dan jumlah variabel prediktor sebanyak 36. Pada model optimal di Kabupaten Lamongan matriks koefisien proyeksi α_{k1} memproyeksikan X_k (variabel GCM) sepanjang 127,03. Matriks koefisien proyeksi α_{k1} secara lengkap dapat dilihat pada Lampiran C.

Model *statistical downscaling* yang optimal untuk meramalkan curah hujan di Kabupaten Bojonegoro, Jember dan Banyuwangi adalah model dengan banyak fungsi $m = 1$. Model-model tersebut disajikan pada Tabel 4.2.

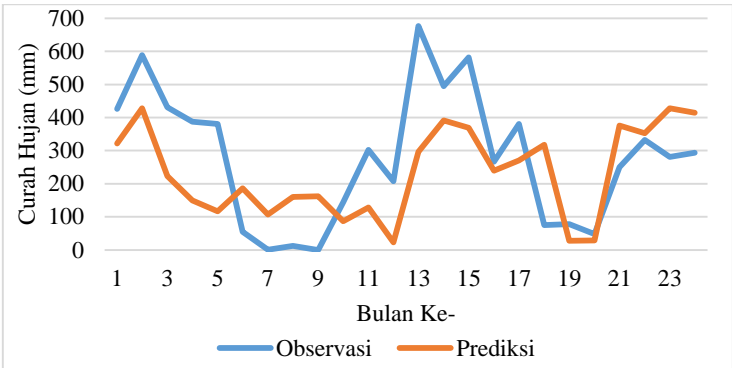
Tabel 4.2 Model PPR Data Curah Hujan di Lima Kabupaten

Kabupaten	Model
Bojonegoro	$\hat{y} = 210,76 + 126,10 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$
Jember	$\hat{y} = 209,38 + 146,53 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$
Banyuwangi	$\hat{y} = 133,60 + 120,80 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$

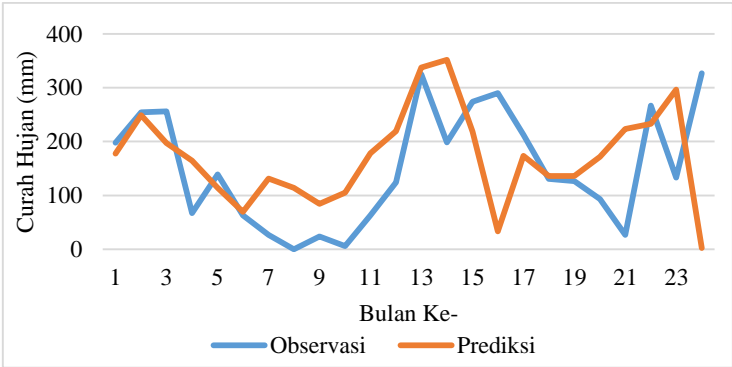
Matriks koefisien proyeksi α_{k1} di Kabupaten Bojonegoro, Jember dan Banyuwangi secara lengkap disajikan pada Lampiran C. Matriks α_{k1} berukuran 1×36 , sesuai dengan banyak fungsi dalam model yaitu $m = 1$ dan jumlah variabel prediktor sebanyak 36. Matriks koefisien proyeksi α_{k1} memproyeksikan X_k sepanjang 126,10 pada model optimal di Kabupaten Bojonegoro. Pada model yang optimal di Kabupaten Jember matriks α_{k1} memproyeksikan X_k sepanjang 146,53. Sedangkan untuk Kabupaten Banyuwangi matriks koefisien proyeksi α_{k1} memproyeksikan X_k sepanjang 126,10 pada model yang optimal.

Perbandingan antara observasi curah hujan dan hasil ramalannya di lima kabupaten yang menjadi sentra produksi padi Jawa Timur berdasarkan data *out-sample* disajikan secara visual pada Gambar 4.6 sampai Gambar 4.10. Gambar 4.6 sampai Gambar 4.10 menunjukkan bahwa hasil ramalan curah hujan di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi mendekati data observasi dan memiliki pola yang sama. Oleh karena itu, model *statistical downscaling* dengan

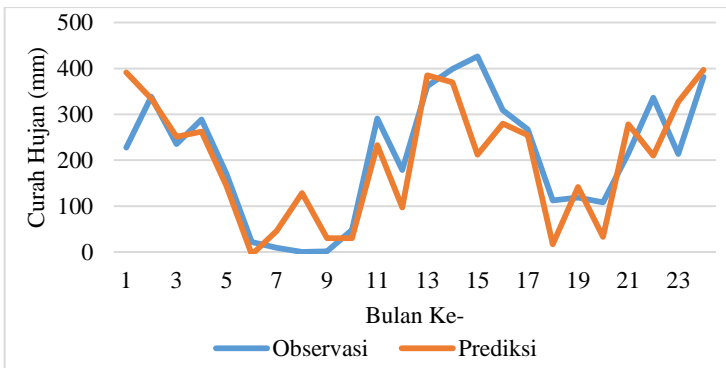
banyak fungsi $m = 1$ merupakan model yang baik untuk meramalkan curah hujan bulanan di lima kabupaten tersebut.



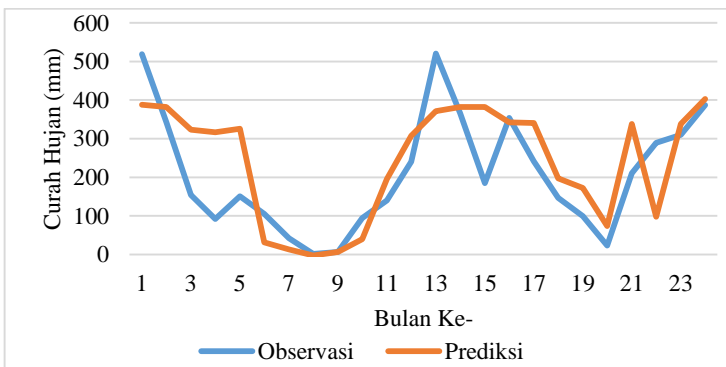
Gambar 4.6 *Plot Antara Observasi dan Ramalan Curah Hujan di Kabupaten Ngawi*



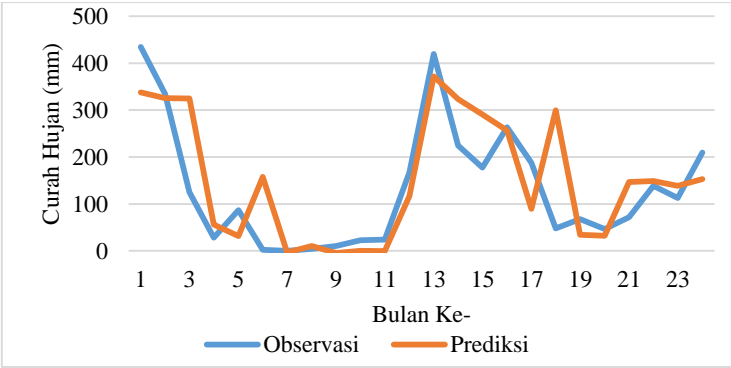
Gambar 4.7 *Plot Antara Observasi dan Ramalan Curah Hujan di Kabupaten Lamongan*



Gambar 4.8 Plot Antara Observasi dan Ramalan Curah Hujan di Kabupaten Bojonegoro



Gambar 4.9 Plot Antara Observasi dan Ramalan Curah Hujan di Kabupaten Jember



Gambar 4.10 *Plot Antara Observasi dan Ramalan Curah Hujan di Kabupaten Banyuwangi*

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan diperoleh beberapa kesimpulan sebagai berikut.

1. Model *statistical downscaling* dengan metode PPR untuk meramalkan curah hujan bulanan di lima kabupaten yang menjadi sentra produksi padi Jawa Timur adalah sebagai berikut.

- Ngawi:

$$\hat{y} = 278,93 + 127,03 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

- Lamongan:

$$\hat{y} = 151,08 + 94,36 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

- Bojonegoro:

$$\hat{y} = 210,76 + 126,10 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

- Jember:

$$\hat{y} = 209,38 + 146,53 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

- Banyuwangi:

$$\hat{y} = 133,60 + 120,80 f_1 \left(\sum_{k=1}^p \alpha_{k1} X_k \right)$$

2. Hasil validasi model *statistical downscaling* dengan kriteria RMSEP menunjukkan bahwa nilai RMSEP terkecil di lima kabupaten sentra produksi padi Jawa Timur diperoleh ketika menggunakan banyak fungsi $m = 1$. Pola antara hasil ramalan dan observasi menunjukkan bahwa hasil ramalan curah hujan di lima kabupaten mendekati data observasi. Sehingga model

dengan banyak fungsi $m = 1$ merupakan model yang baik untuk meramalkan curah hujan di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi.

5.2 Saran

Berikut merupakan beberapa saran yang direkomendasikan untuk pengembangan penelitian selanjutnya.

1. Melakukan analisis peramalan curah hujan dengan beberapa metode sebagai perbandingan, sehingga bisa digunakan untuk menentukan metode terbaik dalam meramalkan curah hujan.
2. Membuat model curah hujan yang berbeda untuk musim hujan dan musim kemarau, karena hal tersebut kemungkinan memiliki hasil ramalan yang baik.

DAFTAR PUSTAKA

- Anitawati, Meika. (2010). Pemodelan Statistical Downscaling Luaran GCM dengan Metode Principal Component Regression (PCR) dan Projection Pursuit Regression (PPR). Tugas Akhir. Surabaya: Institut Teknologi Sepuluh Nopember.
- Fernandez, E. (2005). On The Influence of Predictors Area in Statistical Downscaling of Daily Parameters. *Report Climate* no.09/2005. Oslo: Norwegian Meteorological Institute.
- Friedman, J. H. dan Stuetzle, W. (1981). Projection Pursuit Regression. *Journal of The American Statistical Association*, 376: 817-823.
- [LAPAN] Lembaga Penerbangan dan Antariksa Nasional. (2014a). *Pola Curah Hujan di Indonesia*. Diambil dari website LAPAN: <URL:<http://www.dirgantaralapan.or.id/moklim/edukasi0609pch.html>> pada 15 Februari 2017.
- [LAPAN] Lembaga Penerbangan dan Antariksa Nasional. (2014b). *Dipole Mole (DM)*. Diambil dari website LAPAN: <URL:<http://www.dirgantaralapan.or.id/moklim/edukasi0609dmi.html>> pada 15 Februari 2017.
- Mukarami, T. dan Zadrach, L. D. (2000). Interannual Variability of Convective Intensity Index Over Indonesia and Its relationship with ENSO. *Journal Meteorologi dan Geofisika*, 13: 795-810.
- Ratag, M. A. (2001). Reanalisis Curah Hujan di Indonesia hasil Simulasi Model Area Terbatas Resolusi Tinggi CSIRO DARLAM. Prosiding temu Ilmiah Prediksi Cuaca dan Iklim Nasional. Lembaga Penerbangan dan Antariksa Nasional (LAPAN), Bandung.
- Runtunuwu, E. dan Syahbuddin. (2007). Perubahan Pola Curah Hujan dan Dampaknya Terhadap Periode Masa Tanam. *Jurnal Tanah dan Iklim*, 26:1-12.

- Safitri, Rina. (2012). Model Output Statistics dengan Projectin Pursuit Regression untuk Meramalkan Suhu Minimum, Suhu Maksimum dan Kelembapan. Tugas Akhir. Surabaya: Institut Teknologi Sepuluh Nopember.
- Sosrodarsono, S. dan Takeda, L. (2003). *Hidrologi Untuk Pengairan* (9th ed.). Jakarta: PT. Pradnya Paramita.
- Sucahyono, Dedi dan Ribudiyanto, Kuku. (2013). *Cuaca dan Iklim Ekstrim di Indonesia*. Jakarta: Pusat Penelitian dan Pengembangan BMKG.
- Wigena, Aji Hamim. (2006). Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah hujan Bulanan. Disertasi. Bogor: Institut Pertanian Bogor.
- Wilby, R. L. dan Wigley, T. M. L. (1997). Downscaling General Circulation Model Output: A Riview of Method and Limitations. *Progress in Physical Geography*, 4: 530-548.
- Zorita, E. dan Storch, H. von. (1999). The Analog Method as A Simple Statistical Downscaling Technique: Comparison With More Complicated Methods. *Journal Climatology*, 12: 2474-2489.

Lampiran A. Statistika Deskriptif Curah Hujan Berdasarkan Bulan

1. Kabupaten Ngawi

Bulan	Rata-rata	Standar Deviasi	Koefisien Variasi
Januari	316,20	143,60	45,41
Februari	342,30	106,50	31,10
Maret	305,50	145,90	47,75
April	238,80	85,00	35,59
Mei	125,60	110,40	87,91
Juni	48,80	55,30	113,24
Juli	33,70	49,30	146,30
Agustus	25,25	31,38	124,27
September	37,10	60,10	162,16
Oktober	163,00	128,20	78,62
Nopember	258,40	130,10	50,36
Desember	267,20	110,60	41,38

2. Kabupaten Lamongan

Bulan	Rata-rata	Standar Deviasi	Koefisien Variasi
Januari	250,30	74,70	29,86
Februari	219,80	81,10	36,88
Maret	210,20	64,90	30,90
April	138,80	68,60	49,43
Mei	74,20	58,70	79,08
Juni	51,65	41,84	81,01
Juli	20,68	32,63	157,80
Agustus	10,41	22,26	213,78
September	16,51	24,68	149,48
Oktober	72,40	77,20	106,55
Nopember	153,40	61,20	39,92
Desember	228,40	85,10	37,27

Lampiran A. Statistika Deskriptif Curah Hujan Berdasarkan Bulan (lanjutan)

3. Kabupaten Bojonegoro

Bulan	Rata-rata	Standar Deviasi	Koefisien Variasi
Januari	312,00	102,40	32,81
Februari	286,40	88,70	30,96
Maret	296,30	98,50	33,23
April	222,20	87,20	39,24
Mei	93,10	67,40	72,39
Juni	55,60	55,30	99,47
Juli	29,90	36,15	120,90
Agustus	21,07	30,86	146,44
September	35,30	56,60	160,16
Oktober	118,40	106,30	89,79
Nopember	245,70	119,10	48,48
Desember	283,20	87,40	30,88

4. Kabupaten Jember

Bulan	Rata-rata	Standar Deviasi	Koefisien Variasi
Januari	372,90	121,50	32,57
Februari	351,10	100,70	28,68
Maret	284,60	110,50	38,82
April	188,20	87,70	46,62
Mei	92,30	58,60	63,46
Juni	54,30	54,20	99,74
Juli	28,32	39,58	139,73
Agustus	13,14	17,51	133,30
September	28,50	47,80	167,80
Oktober	118,50	99,20	83,71
Nopember	238,60	87,60	36,69
Desember	345,40	126,90	36,75

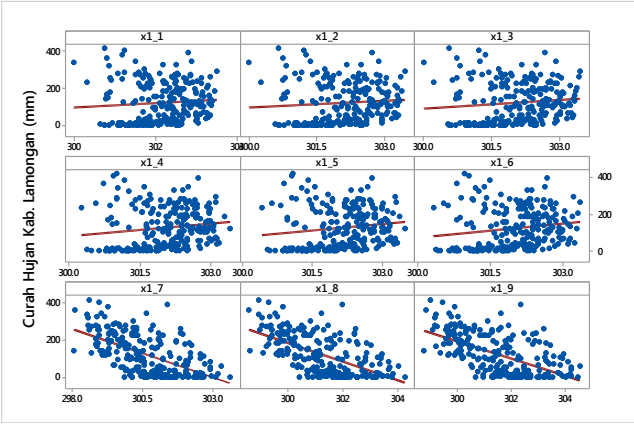
Lampiran A. Statistika Deskriptif Curah Hujan Berdasarkan Bulan (lanjutan)

5. Kabupaten Banyuwangi

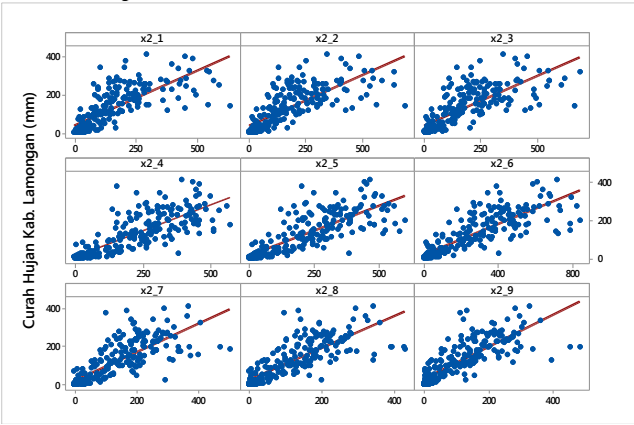
Bulan	Rata-rata	Standar Deviasi	Koefisien Variasi
Januari	323,80	131,60	40,66
Februari	302,70	90,50	29,91
Maret	235,50	109,00	46,27
April	152,40	97,60	64,08
Mei	81,90	64,00	78,21
Juni	51,30	55,00	107,17
Juli	51,50	57,00	110,57
Agustus	26,91	35,28	131,10
September	15,48	24,73	159,82
Oktober	42,50	53,70	126,42
Nopember	90,10	93,60	103,90
Desember	240,90	99,80	41,42

Lampiran B. *Scatterplot* Antara Variabel Respon dan Variabel Prediktor

1. Kabupaten Lamongan
Variabel *Temperatur*

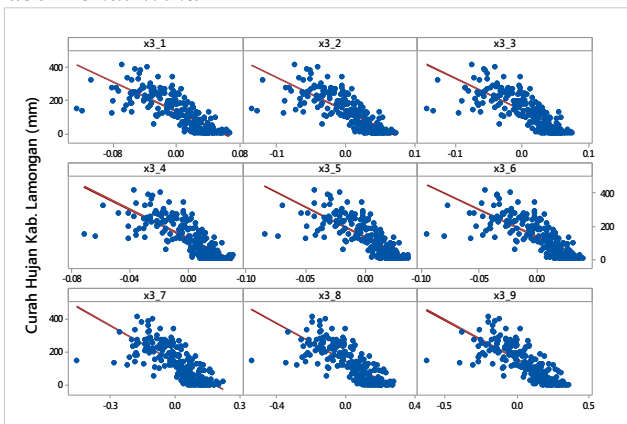


Variabel *Precipitation*

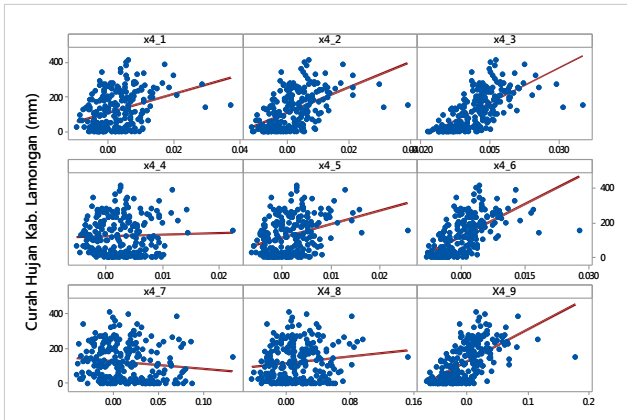


Lampiran B. Scatterplot Antara Variabel Respon dan Variabel Prediktor (lanjutan)

Variabel *Zonal Wind*

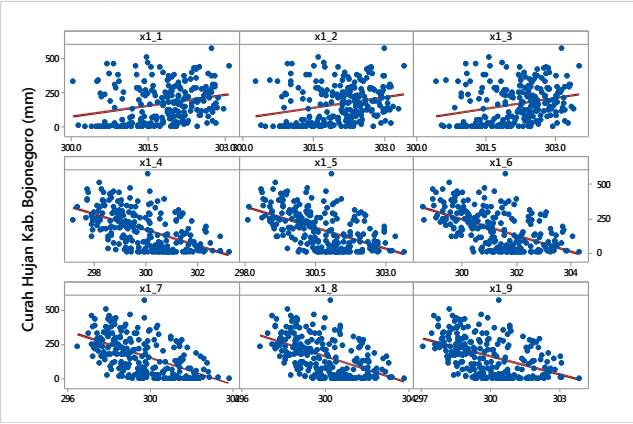


Variabel *Meridional Wind*

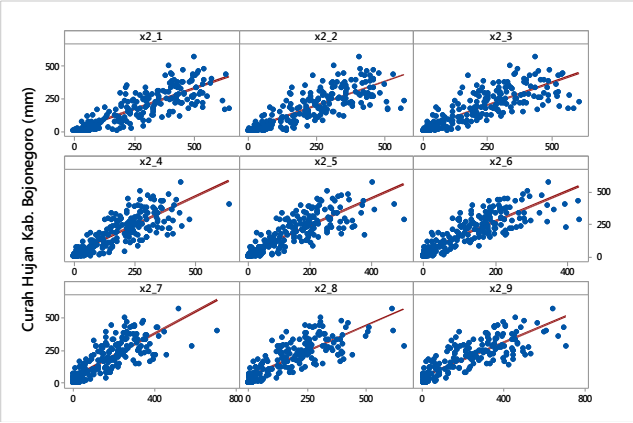


Lampiran B. *Scatterplot* Antara Variabel Respon dan Variabel Prediktor (lanjutan)

2. Kabupaten Bojonegoro
Variabel *Temperatur*

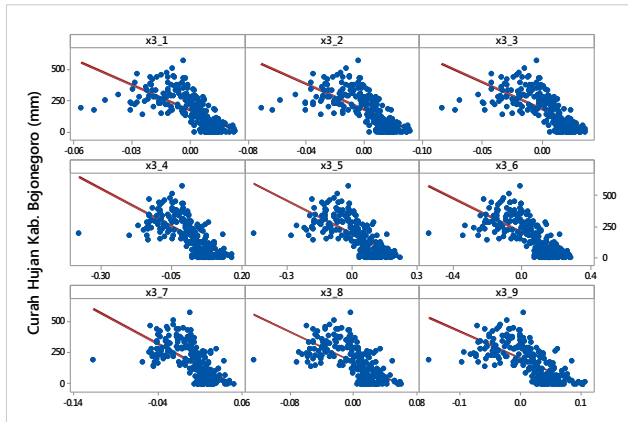


Variabel *Precipitation*

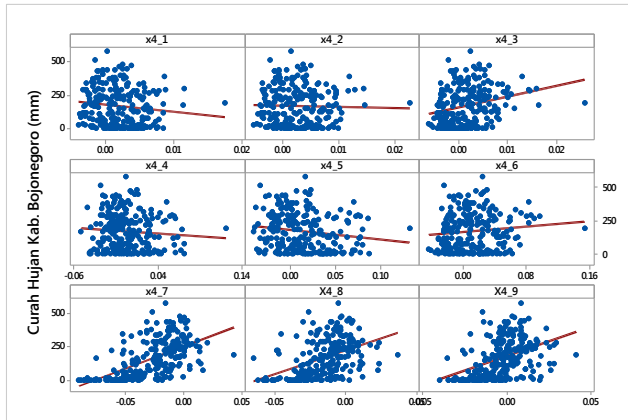


Lampiran B. *Scatterplot* Antara Variabel Respon dan Variabel Prediktor (lanjutan)

Variabel *Zonal Wind*

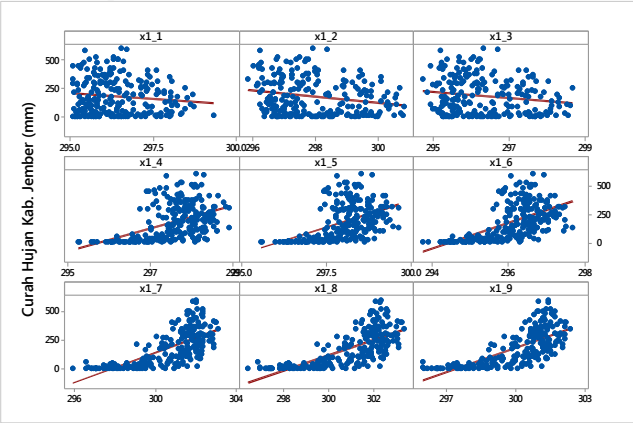


Variabel *Meridional Wind*

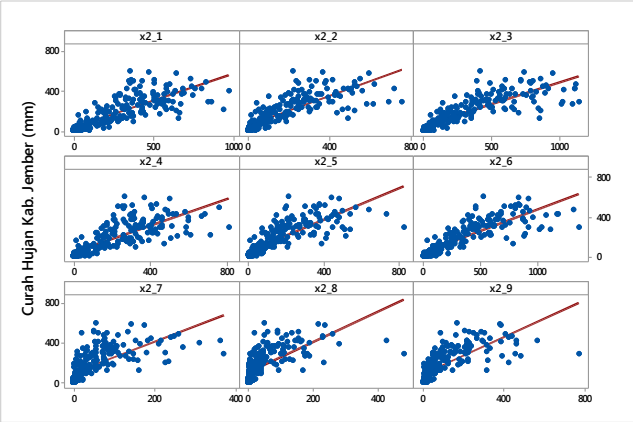


Lampiran B. *Scatterplot* Antara Variabel Respon dan Variabel Prediktor (lanjutan)

3. Kabupaten Jember
Variabel *Temperatur*

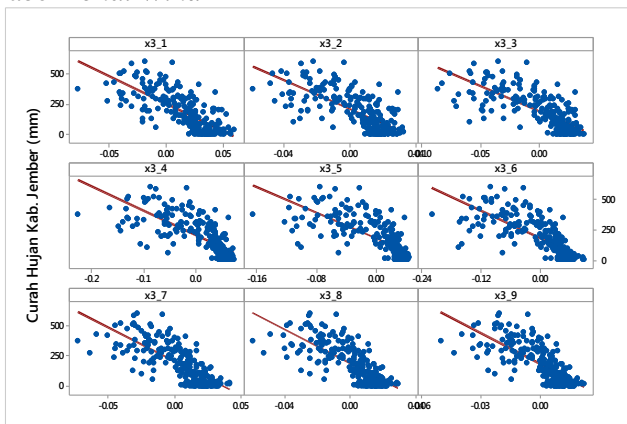


Variabel *Precipitation*

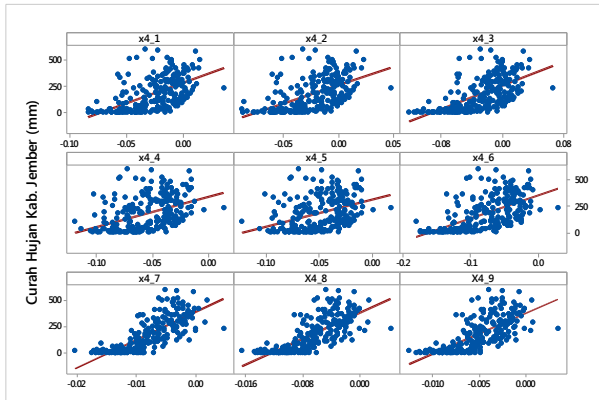


Lampiran B. Scatterplot Antara Variabel Respon dan Variabel Prediktor (lanjutan)

Variabel *Zonal Wind*

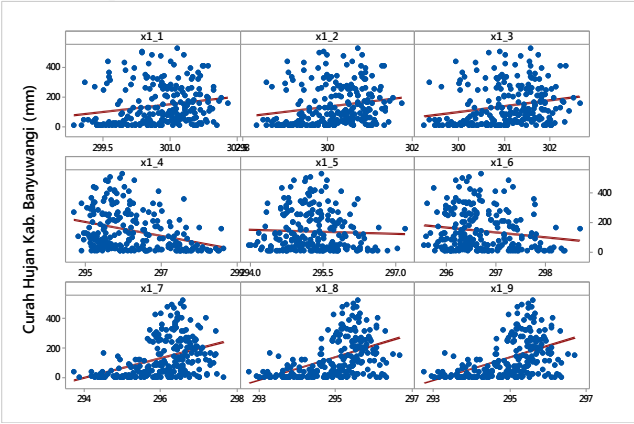


Variabel *Meridional Wind*

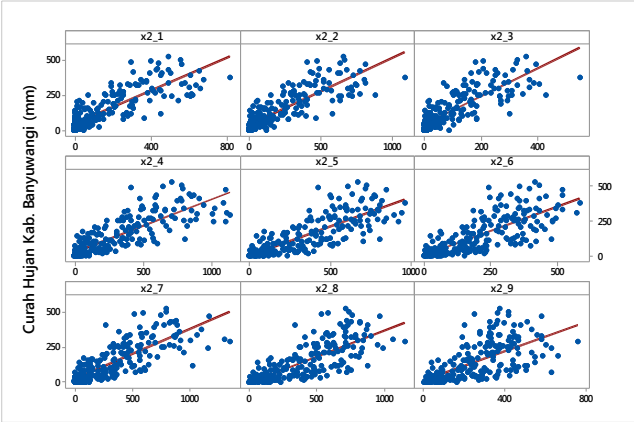


Lampiran B. *Scatterplot* Antara Variabel Respon dan Variabel Prediktor (lanjutan)

4. Kabupaten Banyuwangi
Variabel *Temperatur*

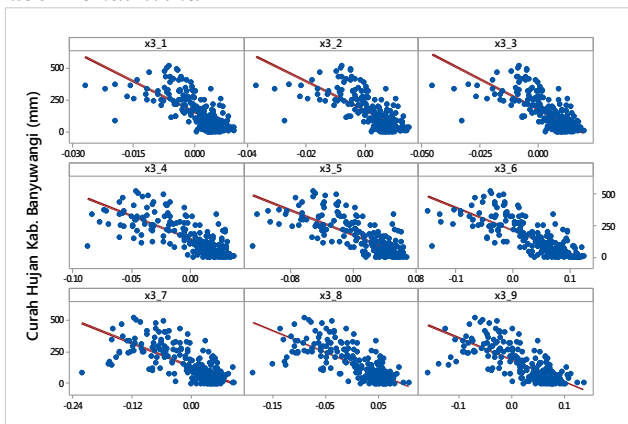


Variabel *Precipitation*

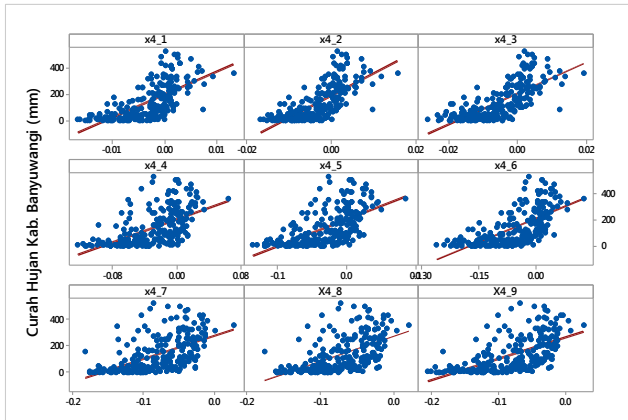


Lampiran B. Scatterplot Antara Variabel Respon dan Variabel Prediktor (lanjutan)

Variabel *Zonal Wind*



Variabel *Meridional Wind*



Lampiran C. Matriks Koefisen Proyeksi α_{k1}

1. Kabupaten Ngawi

$$\alpha_{k1}^T = \begin{bmatrix} -0,03 \\ 0,13 \\ -0,19 \\ -0,10 \\ -0,02 \\ 0,13 \\ 0,33 \\ -0,43 \\ 0,22 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ -0,00 \\ 0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ 0,00 \\ -0,27 \\ 0,03 \\ 0,04 \\ -0,12 \\ -0,03 \\ -0,01 \\ -0,46 \\ -0,32 \\ -0,27 \\ -0,04 \\ -0,19 \\ -0,16 \\ 0,04 \\ -0,04 \\ -0,01 \\ 0,13 \\ 0,05 \\ 0,12 \end{bmatrix}$$

2. Kabupaten Lamongan

$$\alpha_{k1}^T = \begin{bmatrix} -0,15 \\ 0,00 \\ 0,23 \\ 0,18 \\ -0,30 \\ 0,03 \\ -0,02 \\ -0,01 \\ 0,02 \\ 0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ -0,00 \\ -0,30 \\ -0,39 \\ -0,44 \\ -0,08 \\ -0,11 \\ -0,13 \\ 0,17 \\ 0,04 \\ 0,07 \\ -0,08 \\ -0,00 \\ 0,09 \\ -0,05 \\ 0,00 \\ 0,05 \\ -0,38 \\ 0,02 \\ 0,36 \end{bmatrix}$$

$$\alpha_{k1}^T = \begin{bmatrix} 0,05 \\ 0,11 \\ -0,18 \\ 0,11 \\ -0,24 \\ 0,14 \\ -0,01 \\ 0,04 \\ -0,01 \\ 0,00 \\ -0,00 \\ 0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ 0,16 \\ 0,10 \\ -0,02 \\ -0,00 \\ -0,03 \\ -0,23 \\ -0,21 \\ -0,13 \\ -0,09 \\ -0,01 \\ -0,01 \\ 0,42 \\ -0,50 \\ -0,49 \\ 0,11 \\ 0,10 \\ 0,08 \\ 0,06 \end{bmatrix}$$

Lampiran C. Matriks Koefisen Proyeksi α_{k1} (lanjutan)

5. Kabupaten Banyuwangi

$$\alpha_{k1}^T = \begin{bmatrix} -0,19 \\ 0,31 \\ -0,11 \\ 0,01 \\ -0,04 \\ -0,19 \\ 0,12 \\ 0,04 \\ 0,04 \\ 0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ -0,00 \\ 0,00 \\ 0,01 \\ 0,02 \\ -0,07 \\ -0,03 \\ -0,13 \\ 0,14 \\ 0,04 \\ -0,09 \\ -0,03 \\ -0,03 \\ -0,03 \\ -0,40 \\ -0,35 \\ -0,44 \\ -0,38 \\ -0,29 \\ -0,20 \end{bmatrix}$$

Lampiran D. *Syntax Projection Pursuit Regression (PPR)*

```
>Ramalan.Insample<-ppreg(prediktorinsample,responinsample,  
  m)  
>Hasil.Ramalan<-ppreg(prediktorinsample,responinsample,m,  
  xpred=prediktoroutsample)
```

Keterangan:

m = banyak fungsi

Lampiran E. Surat Keterangan Legalisasi Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMIPA ITS.

Nama : Vella Rochmana Asyefa

NRP : 1313 100 073

menyatakan bahwa data yang digunakan dalam Tugas Akhir/~~Tesis~~ ini merupakan data sekunder yang diambil dari penelitian/~~buku~~/~~Tugas Akhir~~/~~Tesis~~/~~publikasi~~ lainnya yaitu:

Sumber : Data Penelitian Dr.Sutikno, S.Si, M.Si dengan Judul “Pengembangan Sistem Pendukung Keputusan untuk Analisis Resiko Produksi Padi dengan Indikator Perubahan Iklim Jangka Panjang dan Jangka Pendek”

Keterangan : 1. Data Curah Hujan Bulanan di Kabupaten Ngawi, Lamongan, Bojonegoro, Jember dan Banyuwangi periode Januari 1991 – Desember 2010
2. *Output* GCM periode Januari 1991 – Desember 2010

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir



(Dr. Sutikno, S.Si, M.Si)
NIP. 19710313 199702 1 001

Surabaya, 13 Juni 2017
Mahasiswa



(Vella Rochmana Asyefa)
NRP. 1313 100 073

*(coret yang tidak perlu)

BIODATA PENULIS



Vella Rochmana Asyeifa, lahir di Blora pada tanggal 27 Januari 1996. Penulis merupakan anak pertama dari dua bersaudara dari pasangan Syaifudin Zuhri dan Sri Setyowati dengan adik laki-laki bernama Nafakhatur Ridwan Asyeifi. Penulis menempuh jenjang pendidikan formal di TK Muslimat NU Japah (1999-2001), SDN Beganjing 2 (2001-2007), SMPN 2 Rembang (2007-2010),

MA Salafiyah Kajen Pati (2010-2013), kemudian penulis diterima sebagai mahasiswa program studi sarjana Departemen Statistika ITS pada tahun 2013 melalui jalur SBMPTN yang terdaftar dengan NRP 1313100073. Pada tahun 2017 penulis berhasil menyelesaikan Tugas Akhir dengan judul “*PEMODELAN STATISTICAL DOWNSCALING DENGAN PROJECTION PURSUIT REGRESSION* UNTUK MERAMALKAN CURAH HUJAN BULANAN DI SENTRA PRODUKSI PADI JAWA TIMUR”. Selama menjadi mahasiswa, penulis tidak hanya aktif di bidang akademik tetapi juga aktif di bidang non akademik. Penulis aktif di Unit Kegiatan Mahasiswa yaitu UKM Cinta Rebana ITS sebagai staf Departemen Event periode 2014-2015 dan Sekretaris Departemen Event periode 2015-2016. Selain itu penulis juga pernah mengikuti beberapa kepanitian kegiatan kampus lainnya. Bagi pembaca yang memiliki saran, kritik atau ingin berdiskusi lebih lanjut tentang Tugas Akhir ini, penulis dengan senang hati dapat dihubungi melalui email vella.rochmana@gmail.com.